The General Equilibrium Incidence of Local Economics Policies

by

Andrew Russell Simon

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2021

Doctoral Committee:

Professor James R. Hines Jr., Chair
Professor John Bound
Assistant Professor Ashley Craig
Associate Professor Kevin Stange

Andrew Russell Simon

arsimon@umich.edu

ORCID iD: 0000-0001-9333-885X

Acknowledgments

This dissertation would not be possible without the help and support of many. I am especially grateful for the members of my dissertation committee: Jim Hines, John Bound, Ashley Craig, and Kevin Stange. I am especially grateful to Jim, as my chair, who provided constant invaluable feedback and guidance on the substance and framing of this work as well as my development as an economist. John Bound and Kevin Stange were not only helpful advisors but also great co-authors who created opportunities for me to learn and explore my interests. Ashley Craig pushed my work in ways, especially theoretically, that improved both the research and my own understanding of the results. My past, current, and future work has greatly benefited from their thoughts, advice, and encouragement.

In addition to my committee, I am grateful to other faculty who importantly shaped my time at Michigan including Charlie Brown, Zach Brown, Susan Dynarski, Ying Fan, Brian Jacob, Joel Slemrod, Jeff Smith, Mel Stephens, and Chris Weiland. My classmates and officemates made every aspect of my time as a graduate student more enjoyable. I would especially like to thank my Chapter 2 co-author Matthew Wilson, and Elizabeth Burland, Jamie Fogel, Fernando Furquim, Max Gross, Thomas Helgerman, Reuben Hurst, Dena Lomonosov, Shawn Martin, Meghan Oster, Stephanie Owen, Shwetha Raghuraman, Michael Ricks, Anna Shaprio, Ellen Stuart, Travis Triggs, and Brittany Vasquez.

My research and graduate training was supported by a grant from the Institute for Education Sciences (R305B150012). This grant provided time and resources for me to conduct my dissertation research and travel to conferences. I received valuable financial support from the Department of Economics, and the Rackham Graduate School. I also thank the

University of Michigan Libraries and Kerby Shedden for helping access to some of the data used in my first chapter.

Finally, I am of course thankful to my partner who provided me with enough cake, and support, to make it through grad school.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

This dissertation contains three essays that examine who benefits and who bears the burden of local government policies, like postsecondary education and the minimum wage. The essays develop general equilibrium structural models to understand the distributional benefits and to assess alternatives.

The first chapter, "Costly Centralization: Evidence from Community College Expansions," examines the implications of public good provision by different levels of government. Higher levels can correct externalities from spillovers, but are often limited to uniform policies and changes in the spatial distribution of public goods and taxes affect where households live. The general equilibrium migration effects can undermine the value of a higher level of government provision because the implied changes in the housing market affect household consumption as well as the aggregate property tax revenue to fund the public good.

I study this trade-off in the context of Texas Community College financing. Leveraging recent increases in the tax base sizes, I find that the communities that joined and their neighbors saw increased aggregate home values and prices after the expansions. This suggests they benefited and more centralization is good on the margin. I then estimate a general equilibrium structural model to measure the optimal level of government provision using Austin Community College as a concrete example. Not all expansions of the tax base are welfare-improving and centralization is not socially optimal because the migration effects are large– general equilibrium housing market distortions reduce the optimal level of government.

The second chapter, "Optimal Minimum Wage Setting in a Federal System" with Matthew Wilson, also measures the relative trade-off of centralized and decentralized government pol-

icy. In a competitive multi-state labor market with two types of workers, increases in one state's minimum wage redistributes to the low-skilled employed from the newly unemployed and high-skilled. A binding policy is optimal if the benefits from redistribution outweigh the costs from migration, which are relatively steeper for local governments. Centralized policy reduces horizontal migration externalities, which improves decentralized minimum wage setting. Our results indicate that decentralized and centralized policy-setting are complements, and the extent to which they are depends on mobility and regional heterogeneity. We then calibrate a model of the continental US and find that joint policy setting leads to a small welfare gain over centralization, and closely resembles the social planner's optimal policies.

The third chapter, "College Choice, Private Options, and the Incidence of Public Investment in Higher Education" with John Bound, studies the general equilibrium effects of state subsidies to public colleges. State funding directly impacts the value of attending the public college since it determines their quality and tuition. Much of the previous work focuses on this channel and implicitly assumes only those who attend the publics benefit. However, since state funding affects the value of the publics, it also affects the relative demand for the privates. Funding decreases increase private college market power, which alters who they admit and the prices they charge. Changes in the college market spill over to the labor market. We calibrate a general equilibrium model of the US higher education system and labor market to illustrate and quantify these mechanisms. Unlike the previous approaches, we find that most of the benefits go to high-income-modest-ability students who only have access to the high quality privates when state spending is high.

Chapter 1

Costly Centralization: Evidence From Community College Expansions

1.1 Introduction

A key function of local governments is to provide public goods and services, which they often finance with property taxes. Prospective residents then choose where to live by trading off public good quality with the associated costs.[1] Tiebout (1956) argues that when communities compete without spillovers, local provision is preferred. However, when the public good spills over to nearby areas, spending is too low and households can free-ride– they benefit, but avoid paying taxes. Higher levels of government can correct the externalities (Oates 1972), but are often limited to uniform policies and changes in the spatial distribution of public goods and taxes affect where households live. The general equilibrium migration effects can undermine the value of a higher level of government provision because the implied changes in the housing market affect household consumption as well as the aggregate property tax revenue to fund the public good.

This paper studies the trade-off between different levels of government provision to fund spatially delineated public goods. In this case, which level of government should set policy? Moving from pure decentralization to pure centralization creates ambiguous welfare effects that importantly depend on the spatial nature of the public good and its spillovers, as well

[1] In 2017, local governments collected $509 billion in property taxes, or 30 percent of their total revenue (Urban Institute 2019), while average household property taxes varied from less than $200 to more than $10,000 at the county level (Cammenga 2019).

as general equilibrium household and housing market responses. Changes in the level of government affect both the level and variance of public goods across space that have important distributional implications for welfare. Although not often considered, intermediate levels of centralization, where a few communities are not part of the tax base but still receive spillovers, can balance the distortions from decentralization or centralization alone.

I address this question in the context of Texas community college financing. Since 83 percent of the land in the state is not part of a tax base, households can live in nearby areas where they still benefit from the public good but do not pay taxes to support it. I first leverage recent expansions of the tax bases across the state, which gradually increased the level of centralization, to test whether they are too small. I then estimate a general equilibrium structural model of household sorting and public good provision to measure the value of different geographic tax base configurations and the optimal level of government provision. These analyses also inform optimal policy for other settings where we expect spillovers to be present, like environmental pollution, as well as place-base policies.

The results indicate that the community college tax bases are too small. Exploiting the staggered timing of recent expansions, I find that both home prices and aggregate home values rise in the newly annexed and other neighboring communities. The net benefits from the change in public goods and taxes are capitalized into home prices, while values additionally capture the effects from household sorting. Increases imply the communities are better off and that increased centralization is welfare-improving on the margin.[2] However, large changes in the boundaries may not have the same positive effects if the benefits are spatially concentrated or if household sorting responses create negative externalities.

I then estimate a general equilibrium structural model to measure the optimal level of government provision using Austin Community College as a concrete example. Since only 17 percent of the land in the state is currently part of a tax base, the previous results do not

[2] Aggregate home values and prices provide information about the overall attractiveness of a community and its resources. Increases in both therefore reflect benefits to the community as a whole, but do not imply that all households living there are better off.

indicate how much they should expand. Using information on household location decisions, housing expenditure, and home sales, the model informs how households respond to changes in public goods and taxes. Combining the model with the policy variation from the recent tax base expansions, I then decompose the public good contribution of community college expenditure across space. I find that the returns to spending are large near the campus regardless of whether a community is in the tax base but decrease with distance, and decrease slightly more sharply for those not in the tax base.

The additional public good benefits from joining the tax base are an inverted-U shape in the distance from the campus. Residents in communities close to the college can more easily access it, even when not part of the tax base. Joining therefore leads to few new services and redistribution to the original tax base. The value of joining increases with distance, up to a point. Communities farther away have less easy access when not part of the tax base, but the expansion stretches the geographic scope of the services. However, the four farthest communities only receive a new tax burden after joining.

Austin Community College centralization, which increases the number of communities in the tax base from nine to 26, raises average per capita welfare by \$29 over the current boundary configuration, or about \$23 Million for the metro area. This welfare gain is equivalent to 15 percent of the college's yearly tax revenue. The benefits are concentrated in the newly annexed areas. All but six of communities that join receive additional public goods that are worth the new tax burden. The implied increased value of living in those areas leads to in-migration and the resulting change in prices from added congestion pushes almost all of the gains to the original landowners. With more property value to fund the public good, the statutory tax rate falls and the original tax base also benefits, but this largely acts as redistribution towards the high-income close to the college.

Household sorting from the change in the distribution of public goods and taxes across space also imposes negative housing market externalities on neighboring college districts in the metro area. Those areas become relatively less attractive and residents leave. Landowners, therefore, bear most of the burden. The distributional changes in prices and public goods also leads some low-income households to lose access to inexpensive communities, making them worse off as well; centralization reduces community heterogeneity in public goods and taxes, which stops some households from finding a good match for their preferences. The state, as an even higher level of government, could play a role to correct these externalities. However, Texas cannot legally levy income or property taxes, and so the maximum possible level of centralization in this case is limited.

However, not all expansions of the tax base are welfare-improving and centralization is not socially optimal. Large changes in the boundaries create negative externalities from household sorting. On average, the changes impose negative pricing effects on poor households who dislike high rental costs, but benefit the high income near the college and landowners. While many expansions that increase the level of public good spending are worth the added congestion, centralization goes too far. In fact, there are nearly 150 tax base configurations with higher social welfare than centralization. For example, an Austin Community College tax base that covers all but one community (school district) raises per capita welfare by an additional $222 over centralization. The general equilibrium sorting externalities play an important role and a partial equilibrium approach that does not account for them would expand the tax base more than is optimal.

The housing market externalities from expanding geographic coverage of the tax base can be so large that adding new areas lowers the total property value to fund the public good. Some expansions therefore lead to higher statutory tax rates while simultaneously imposing an additional tax burden on more communities. Although a large tax base helps reduce free-riding, general equilibrium responses can create unintended consequences that undermine the value of centralization. When the extra revenue finances additional spending or public good quality, the negative consequences are exacerbated.

The main contribution of this paper is quantifying spillovers, and the relative merits centralization and other levels of government public good provision. To do so, I build on the fiscal competition and local public goods models (Epple and Sieg 1999).[3] I extend the previous work in two ways. First, my model allows for inter-community spillovers, which are an important factor for determining the relative value of centralization, as highlighted in Oates' Decentralization Theorem (1972).[4] Second, I combine information on household sorting behavior with policy variation over time to estimate the public good production function. This type of locational sorting model has been used to study pollution (Sieg et al. 2004; Smith et al. 2004), private school vouchers (Ferreyra 2007), and open space protections (Walsh 2007). The previous work using this framework considers a cross-section of communities and relies on functions of the income rank as instruments to estimate the value of amenities. My approach combines causal inference with a general equilibrium sorting model, similar to Bayer et al. (2007).

This paper also adds to the empirical literature on public good spillovers and centralization, of which there is relatively little on either or their interaction.[5] Solé-Ollé (2006) in Spain and Jannin and Sotura (2019) in France find public good spillovers are large, similar to the findings in this paper, but do not measure the value of different levels of government provision, while Calabrese et al. (2012) find that centralization is often more efficient, but assume that there are no spillovers. Closest to this paper, Banzhaf and Chupp (2012) estimate the value of centralized environmental policy with spillovers. In my setting, spillovers from production additionally create externalities from their impact on household location decisions. These effects interact since migration changes the size of the tax base and therefore the per household cost of providing the public good, which further affects household residencies.[6] This paper also considers intermediate cases between decentralized policies and

[3]Kuminoff et al. (2013) and Holmes and Sieg (2017) provide an overview of this methodology and recent advances.

[4]More specifically, "in the absence of cost-savings from the centralized provision of a [local public] good and of interjurisdictional externalities, the level of welfare will always be at least as high (and typically higher) if Pareto-efficient levels of consumption are provided in each jurisdiction than if any single, uniform level of consumption is maintained across all jurisdictions" (Oates 1972).

[5]However, there is a large theoretical literature beginning with Oates (1972) and Bewley (1981). Recent examples include Cheikbossian (2008), and Bloch and Zenginobuz (2006). Besely and Coate (2003) study a political economy approach to this problem.

[6]Simon and Wilson (Forthcoming) also study the effect of migration on the relative value of policy setting

centralization to understand the optimal level of federalism.[7]

1.2 Community College Taxing Districts in Texas

The Texas government divided the state into 50 community college service areas. Each service area's mission is to provide postsecondary education to the entire region, but is only able to levy an ad valorem property tax on a subset, the taxing district. While the service areas cover almost all of Texas, the taxing districts do not. If spillovers are positive, households can sort to avoid property taxes but still receive some benefits. Increasing the size of the tax base should limit free-riding. However, if the public good is concentrated near the campus, expanding the tax base under the state required uniform tax rate would put an undue burden on communities far away. This system where households can free-ride from community college services by living outside the taxing district is common in the United States and is also used in Arkansas, Colorado, Idaho, Kansas, Michigan, Missouri, Montana, New Mexico, Ohio, Oregon, and Wyoming.[8]

Since 1990, 49 communities in Texas, usually independent school districts (ISDs), across seven service areas voted to join their taxing district. Groups of ISDs form coalitions to provide the public good to the service area. As the coalitions grow, the level of centralization increases and these expansions provide a test of its value. It is illegal for the state of Texas to levy its own property or income taxes, and so the relevant units of governments to set policy in this setting are ISDs and service areas, limiting the level of centralization.

The public good benefits of community colleges include having a more highly educated workforce, less expenditure on social services, increased charitable giving, greater involvement in civic life, improved social cohesion, and a reduced crime rate (Mellow and Heelan

by different levels of government in the context of US minimum wages.

[7]The focus here is on optimal geographic size of the government to provide local public goods, as in Saarimaa and Tukiainen (2014), Weese (2015), and Tricaud (2019) who study jurisdiction mergers and find support for similar mechanisms related to distance. There is also a related literature that studies which public goods should be provided by local or central government; see Bednar (2011) for an overview and a discussion of other federalism issues in political science.

[8]Parts of Illinois were not part of a taxing district until 1990. Additionally, community colleges in 29 states rely at least partially on local appropriations, which would still allow households to partially free-ride by living in low-tax community college regions and benefit from their neighbors.

2014; McFarlin et al. 2017). A more educated society benefits not just those who receive the education but has a positive externality on the other residents, providing social returns in excess of the aggregate private returns.[9] One potential concern of increasing community college access for generating the public good is that it may divert students away from bachelor degree programs and reduce the aggregate amount of human capital, even though it has significant value-added for marginal students (Mountjoy 2019). However, Denning (2017) and Action (2019) find that residency in a community college taxing district, which increases access through reduced tuition, leads to more two and four year college enrollment and graduation, and therefore unambiguously higher levels of aggregate human capital.

The community college also strengthens other local amenities. For example, 10 percent of high school students were enrolled in dual credit courses in 2018, increasing the value and quality of K-12 education. The state hopes that by 2030, at least 30 percent of high school students will complete at least 12 dual credit hours (THECB 2019). Community colleges also strengthen community infrastructure by training many first responders, like police and firefighters, and health care providers (Austin Community College 2020). These benefits are not just for those in the taxing district; for example, when a college trains police, those officers get jobs throughout the surrounding area, not just in the tax base.

For communities to receive these benefits from having a community college, it needs to be appropriately funded. A larger tax base gives the college more, and more stable access to funding. This has become increasingly important over time as state appropriations continue to fall. Over the past 30 years, state appropriations for Texas community colleges fell from 66 to only 23 percent of operating revenue, making colleges more reliant on local property taxes (TACC 2019). Before voting to join the taxing district, the college and community to be annexed agree on a service plan, which outlines how the college's revenue will be used.[10] A goal of the annexation is to expand the geographic scope of the benefits.

[9] Moretti (2004) provides a summary of theoretical and empirical literature on the social returns to human capital. This includes the effects on crime and voting but also positive productivity spillovers, where increases in the aggregate human capital in a city increases the wages of all workers.

[10] See Appendix A.1 for an example.

The focus of this paper is on financing the public good, but community colleges provide a private good as well for those who attend. Taxing district residents face lower tuition and so they attend more frequently (Denning 2017; Martorell et al. 2018; Action 2019), but households in other parts of the service area and state also attend. Given the private benefits, why not finance community college expenditure with tuition alone or increase non-resident tuition to compensate for the spillovers? This would instead impose free-riding along the attendance margin. Households who do not pay tuition because they are not enrolled would benefit from the public good without contributing to its financing. If the college relied solely on tuition, police officers, for example, would have to pay more for their own training, likely leading to an under provision of their services to help lower the crime rate.

1.3 Are Community College Tax Bases Too Small?

If the benefits of having a community college spill over to nearby areas outside the taxing district, households can free-ride and therefore the tax base would likely be too small. The extent to which households can free-ride depends on the size of the spillovers relative to benefits in the taxing district and how these change with distance from the college. If the benefits decrease sharply with distance, even for those in the taxing district, then communities may not want to join. The state required uniform rate may also put too large of a burden on communities far from the campus, regardless of the size of the spillovers.[11]

I first estimate the effect of joining the tax base on the newly annexed school districts (ISDs) using a differences-in-differences framework to test whether the current tax bases are too small.[12] If the value of the new services from joining the the tax base outweigh the additional taxes owed, the value of property in the area should increase since the public good is capitalized into home prices and more households will want to move in. Increases in prices and aggregate home values therefore provide a measures of the net benefits. Annexations

[11] The taxing district may also be too small because of political frictions. The state requires that the tax base is contiguous, which limits the number of communities that are eligible to be annexed.

[12] While the majority of annexations are of ISDs, when looking at the ISD level data, partial annexations are treated as if the ISD were completely annexed. This is not a problem when looking at finer levels of geography, like Census Tracts.

may also impact the original tax base and the other ISDs in the service area through how they affect the public good and household sorting. I provide suggestive evidence that the current tax bases are too small since average home prices and aggregate values weakly increase in all communities after recent expansions– more centralization on the margin is welfare-enhancing.

1.3.1 Data

The Texas Independent School District Self-Reports of Value from the state comptroller provides information on aggregate property values. Beginning in 1991, each school district reports the total property value within its borders by type. This data forms a balanced panel from 1991 to 2016. To capture the effect on home prices, I use Census Tract level "repeat-transactions" home price indices (HPI) developed in Bogin et al. (2019) and provided by the Federal Housing Finance Agency (FHFA). Census Tracts are matched to ISDs using the correspondence tables from the Missouri Census Data Center.

To measure the distributional effects of increased centralization across space, the analysis divides school districts into three types of communities: (1) those in the taxing district by 1991; (2) those that joined after 1991; and (3) those that have not yet joined. Table 1.1 presents summary statistics by community type from the 2016 Self-Reports of Value data as well as the 2014-2018 American Community Survey from IPUMS (Manson et al. 2019). The current taxing districts are mainly urban areas that only cover about 17 percent of the land in the state, but 71 percent of the population. They also tend to be more highly educated, and have slightly higher incomes, home values and rents. Since the taxing districts must be contiguous, communities in the tax base are closer to the main campus. Additional annexations add a tax burden to lower income, far away communities.

1.3.2 Impacts of Recent Annexations

To measure the effect of joining a taxing district, the estimation strategy leverages the staggered timing of annexations and estimates a difference-in-differences using the following event-study specification on the set of communities that voted to join between 1995 to 2011:[13]

$$\log(y_{jt}) = \sum_{k=-4}^{5} \beta_k \mathbb{1}(j \in \text{TD})_t + \phi_j + \xi_t + \epsilon_{jt} \tag{1.1}$$

where y_{it} is the outcome of interest, for example, aggregate home values or home price indices, ϕ_j are geographic fixed effects at the same level of the data, which are either school districts or Census Tracts, and ξ_t are year fixed effects. The range of k is chosen so that the effects are estimated using a balanced sample of ISDs.

Aggregate home values and price indices provide complementary information on the benefits to a community. Values reflect the size of the total tax base including the effects on migration and housing supply that can be used for public goods and services, while prices reflect the gains to existing home owners.[14] β_k estimates the effect of joining the tax base k years after the vote with β_0, or the estimated effect for the year of the vote, normalized to 0. If households value joining more than the added tax, β_k will be positive for k greater than 0, that is, the net benefit is capitalized into home prices.[15] Although the tax burden begins at year 1, β_1 does not necessarily capture the difference in realized benefits. Prices may adjust before the community begins to benefit if households are forward looking.

The main identifying assumption to measure the causal effect of joining the taxing district is that in the absence of treatment, treated and comparison communities' outcomes would evolve similarly. Although this assumption is not directly testable, Equation 1.1 estimates the relative differences in the few years prior to annexation. If the coefficients for $\beta_k,\ k < 0$, are zero, then this provides suggestive evidence that the parallel trends assumption holds. Since all communities in the sample are eventually treated, the comparison communities are

[13] The date range is picked so that the event-study coefficients can be estimated for several years before and after the vote for each ISD in the sample. Annexations in 1991 and 2016 are therefore excluded, although they take place during the same period.

[14] High home prices have an ambiguous welfare effect on individual households conditional on public expenditure. It becomes more costly to live in an area for both owners and renters but increases housing rents for owners. The policy counterfactuals in Sections 1.6 and 1.7 decompose these channels.

[15] Ross and Yinger (1999) discuss the economic intuition behind home price capitalization and review the literature to asses public expenditure capitalization (e.g. Oates 1969 and Brueckner 1982), while Baum-Snow and Ferreira (2015) provide an overview of causal inference methods estimate similar hedonic models.

those that will be annexed in the future and this design leverages the staggered timing of votes.

The parallel trends assumption may fail if voters are only willing to join the tax base during periods of high local economic growth, conditional on year fixed effects. Alternatively, a college may only try to organize an annexation in harsh economic conditions when it needs additional revenue. Since state appropriations have been falling for over 30 years, business cycle fluctuations likely do not affect the college's yearly incentives.[16] Anecdotally, the ballot measures are costly to organize and the exact timing of when they will reach the ballot is potentially quasi-random. For example, Houston hoped to annex Katy, North Forest, and Spring Branch ISDs in 2004 (Bricker 2008) but only North Forest and Spring Branch voted in 2009. At the opposite extreme, Pflugerville appears to have arranged for the proposition to be on the ballot in less than one year.[17] The event-study specification additionally helps overcome some identification issues of two-way-fixed-effect difference-in-differences with differential timing (Goodman-Bacon 2019).

Panel (a) of Figure 1.1 plots the estimated coefficients for the effect on home values in solid black circles. In the few years before the vote, the point estimates are very close to zero, although they are somewhat imprecise. Denning (2017) provides additional suggestive evidence that the timing of the annexations are quasi-random using student level data. He finds no evidence of pre-trends on community college and four-year enrollment or graduation rates, and the annexation is unrelated to student characteristics including gender, race, economic disadvantage status, and limited English proficiency. He also finds that simply holding a vote does not impact community college take-up by looking at failed annexations, which suggests the timing is not correlated with other simultaneous changes in a community's human capital.

The year after the vote passes, aggregate values remain constant. Under the new tax

[16] There must also not be any other contemporaneous events, for example due to other propositions or elections. One partially reassuring note is that the majority do not take place during an election year.

[17] The entire process may have been longer since it is unclear from news reports on the subject how long the community planned before seeking resident signatures.

burden, households do not appear to be immediately worse off. Over the next few years, home values steadily increase before leveling off, which suggests that the benefits accumulate over time as more households have access to the public good. The point estimates imply that aggregate home values increase about 2.5 percent from joining the tax base on average. We see a very similar pattern from the solid black circle in Panel (b) for the effect on home prices, which increase by about 1 percent after the vote on average, although they are imprecisely estimated. Given the relatively little policy variation, it is not surprising that the estimates are noisy, but the home valuation and price data come from different sources on different sets of houses in each ISD and together show a very similar pattern. The noise may also capture important spatial heterogeneity in the benefits from joining the taxing district. For example, some communities far from the campus could have very little access to the college even after joining but would be worse off from the tax burden. In the structural model in Section 1.5, I model the heterogeneity across space.[18]

The estimates from Equation 1.1 measure the unconditional effect of joining the tax base for a community. However, after an annexation, there is an ambiguous effect on other communities' decisions to join. First, it may increase the incentive to free-ride. With a larger tax base, the college has additional revenue to finance its operations that benefit both the taxing district and areas receiving spillovers. The college may also build a campus in the newly annexed community, stretching the college's reach. Alternatively, since increases in the tax base reduce the financial burden for any single community, the cost of joining decreases.[19] To better capture the partial equilibrium effect of the annexation, the open square estimates plotted in the figure control for the log of the number of other school districts in the tax base, but the point estimates do not change much.[20]

[18]The results are consistent with Denning's (2017) results on attendance, which links the effects shown in Figure 1.1 to the mechanisms described in the previous section; there are positive social returns to additional education and these are capitalized into home prices and values. Some of the estimated effect is likely due to private returns as well.

[19]Di Porto and Paty (2018) find that when communities join an inter-municipal jurisdiction, it induces others to as well in France.

[20]Appendix A.3 presents the estimated coefficients on the number of residential properties, which also increase after the annexation. The appendix also estimates the impact on K-12 education to show the

The annexation may also affect the other communities in the service area. With the additional tax revenue, the initial tax base will see some combination of increased expenditure and a decreased tax rate. Increased expenditure should benefit the communities receiving spillovers while also making them more attractive for households to free-ride. Equation 1.2 estimates the effect of the size of the tax base on the initial taxing district and the remaining parts of the service area:

$$\log(y_{jt}) = \gamma \log(\# \text{ in TD}) + \phi_j + \xi_t + \epsilon_{jt} \tag{1.2}$$

where γ is the elasticity of y with respect to the number of school districts in the tax base. γ is identified based on the staggered timing of successful annexations and # in TD is arguably exogenous because the initial tax base and remaining parts of the service area do not control whether a community is annexed. The results in Table 1.2 provide further suggestive evidence that the current tax bases are too small; all communities in the service area appear to weakly benefit from a larger tax base and increased centralization.

1.3.3 Impacts on the Community College Market

How does the size of the tax base affect the quality of the public good provided? With access to more revenue, the college can spend more on instruction to improve quality, increase the geographic scope of its services, decrease the statutory rate to lower the tax burden, or some combination of all three. Using a similar expression to Equation 1.2 at the community college level, Table 1.3 presents the estimated effect of expansions in the tax base, measured by the aggregate market value of property in the taxing district, on the tax rate and instructional expenditure per student using all 50 colleges in the state.

Column (1) shows that increases in the size of the tax base lead to a lower statutory rate. With more revenue, the college is able to decrease the tax burden on each individual. However, the magnitude is quite small. The average tax rate over the sample period was

---

impact on other public goods. The results suggest that the K-12 spending per student are unaffected. Appendix A.4 also suggests that there is no impact of annexation on aggregate firm property values.

about 16 cents for every $100 in property value, or 0.16 percent. For the average household in a taxing district, a 0.2 percent reduction in the rate is about $5 of tax savings each year on a base of about $230. The decrease in the tax rate especially benefits households with larger homes, who tend to be high income. The change may therefore be regressive. Since the decrease in the rate is small, the college collects more tax revenue.

However, Column (2) shows no increase in instructional expenditure per student. Instead of spending more per student, the college is likely spending its extra revenue in additional areas as described in Section 1.2. Changes in the size of the tax base may not change college quality but rather the geographic scope of who benefits. Increases in the size of the tax base may also not lead to significantly more spending because, on average, expansions include more lower income households far from the college. In that case, the new median voter after the annexation has weaker preferences for per capita expenditure conditional on the cost of providing the public good.

1.4 Household Sorting and Public Good Provision

The previous results provide suggestive evidence that community college tax bases are too small, but how much should they expand? Or, what is the optimal level of government provision? This section describes a general equilibrium model of household sorting and public good provision with spillovers to understand the consequences of future expansions and increased centralization.

The paper builds on the previous structural models of fiscal competition and public good provision, beginning with Epple and Sieg (1999). This framework estimates how households respond to changes in public good provision and taxes from the household location decisions and housing prices in the data. Then, based on the model's parameters and policy variation from recent annexations together, I decompose the public good into its components, including the spatially delineated benefits of community colleges. This paper uses causal inference methods in the context of the model to estimate the public good production function. The previous models in this framework use only a cross-section and therefore rely on instruments, usually functions of the income rank of the community.[21]

1.4.1 Household Behavior

The economy has a continuum of households in a metropolitan area divided into J communities with fixed boundaries. Each community j has a local housing market, and provides a public good g_j financed by ad valorem property tax t_j. Households are differentiated by income y and a preference for the public good α. They freely search for the community that best suits their needs trading-off the value of the public good and the tax inclusive cost of housing services, $p_j = (1 + t_j)p_j^h$. (α, y) are jointly log-normally distributed and households' indirect utility function is given by:

$$V(g, p, y, \alpha) = \left\{ \alpha g^\rho + \left[\exp\left(\frac{y^{1-\nu} - 1}{1 - \nu} \right) \exp\left(- \frac{Bp^{\eta+1} - 1}{1 + \eta} \right) \right]^\rho \right\}^{\frac{1}{\rho}} \tag{1.3}$$

which implies housing demand by Roy's identity

$$h = Bp^\eta y^\nu \tag{1.4}$$

where B is a housing demand shifter, ρ is the elasticity of substitution between the public good and private consumption, and η and ν are the price and income elasticities of housing, respectively. If $\rho < 0$, then the single-crossing property holds, which is key to characterizing the equilibria.[22]

Since households choose the community j that maximizes their indirectly function, the

[21] Kuminoff et al. (2013) reviews the other types of equilibrium sorting models, like the random utility household sorting models, as in Bayer et al. (2004).

[22] Under the model's parameterization, the slope of the indirect indifference curve in the (g, p) plane is:

$$\frac{\alpha g^{\rho-1}}{Bp^\eta} \left[\exp((y^{1-\nu} - 1)/(1 - \nu)) \right]^{-\rho} \left[\exp((Bp^{\eta+1} - 1)/(1 + \eta)) \right]^{-\rho}$$

which is monotonic in y for a given α, and vice versa, as long as $\rho < 0$. This is empirically tested later.

set of households in each community is:

$$C_j = \{(\alpha, y) | V(\alpha, g_j, p_j, y) \geq \max_{i \neq j} V(\alpha, g_i, p_i, y)\} \tag{1.5}$$

Epple and Platt (1998) and Epple and Sieg (1999) show that under the model's assumptions in an locational equilibrium, there must be an ordering of community pairs, $\{(g_1, p_1), ..., (g_J, p_J)\}$, such that there will be a set of households that are indifferent between two adjacent in median income rank communities

$$I_j = \{(\alpha, y) | V(\alpha, g_j, p_j, y) = V(\alpha, g_{j+1}, p_{j+1}, y)\} \tag{1.6}$$

The boundary indifference condition between j and $j+1$ is

$$\ln(\alpha) - \rho\left(\frac{y^{1-\nu} - 1}{1-\nu}\right) = \ln\left(\frac{Q_{j+1} - Q_j}{g_j^{\rho} - g_{j+1}^{\rho}}\right) \equiv K_j \tag{1.7}$$

where $Q_j = \exp\left(-\frac{\rho}{1+\eta}(Bp_j^{\eta+1} - 1)\right)$. The sorting equilibrium therefore relates household location decisions and prices to the public good.

The equilibrium has two additional properties that are useful for estimation– stratification and increasing bundles. Let $y_j(\alpha)$ be the implicit function defined by Equation 1.6. Then stratification implies that for each α, households in community j have income y given by $y_{j-1}(\alpha) < y < y_j(\alpha)$. Increasing bundles states that for two communities i and j with $p_i > p_j$, then $g_i > g_j$ if and only if $y_i(\alpha) > y_j(\alpha)$.

1.4.2 Communities

Aggregating household decisions from the previous subsection characterizes communities. Equation 1.5 implies that the percent of households in, or market share of, community j is

$$n_j = P(C_j) = \int_{-\infty}^{\infty} \int_{K_{j-1}+\rho\frac{y^{1-\nu}-1}{1-\nu}}^{K_j+\rho\frac{y^{1-\nu}-1}{1-\nu}} f(\ln\alpha, \ln y) d\ln\alpha d\ln y \tag{1.8}$$

where the functional form of the bounds of integration are determined by the choice of the indirect utility function. Aggregate housing demand for j can similarly be found by integrating over Equation 1.4. Housing supply is assumed to be given by

$$H_j^s = l_j(p_j^h)^\tau \tag{1.9}$$

where l_j is a measure of the amount of land in the community, and τ is the price elasticity of supply.

The value of housing in j determines the size of the community's property tax base to fund the public good. j's budget constraint is therefore $t_j p_j^h H_j/n_j = c_j(g_j, s)$, where $c(\cdot)$ is the cost per household of providing the public good in state of the world s. In my application, s indexes the communities in Austin Community College's tax base.[23] The budget, as a function of prices and aggregate housing values in the community, and the housing market clearing condition implicitly define the set of feasible policies or the government-services possibility frontier (GPF). Within each community, (g^*, p^*) is a majority rule equilibrium if there is no other point on the GPF that would beat it in a pairwise vote. Voters, therefore, maximize their indirectly utility function subject to this feasibility constraint. Epple et al. (2001) show that, under single crossing, there is a locus of pivotal voters in each community:

$$\int_{-\infty}^{\infty} \int_{K_{j-1}+\rho\frac{y^{1-\nu}-1}{1-\nu}}^{L_j+\rho\frac{y^{1-\nu}-1}{1-\nu}} f(\log\alpha, \log y) d\ln\alpha d\ln y = \frac{1}{2}n_j \tag{1.10}$$

$$\text{with} \quad L_j \equiv \ln(\alpha) - \rho\left(\frac{y^{1-\nu}-1}{1-\nu}\right) = \ln\left(\frac{BQ_j p_j^\eta p_j'(g)}{g_j^{\rho-1}}\right) \tag{1.11}$$

Similar to Equation 1.7 from the sorting equilibrium, these equations relate household location decisions and prices to the public good level in each community. If voters are myopic, that is, if they believe that the distribution of households across communities is not affected

[23]The cost of providing the public good is community specific because it depends on the amount of exogenous amenities in that location, like living near a river. It is also a function of the state of the world because the size of the tax base determines the amount of spending and the tax rate. The extent to which these change with the level of centralization depends on returns to scale for providing the services.

by changes in the public good, then

$$\frac{\partial p_j}{\partial g_j} = \frac{\tau(1+t_j)}{(\tau-\eta)+t_j\tau(\eta+1)}\frac{c'(g_j)}{H_j/n_j} \tag{1.12}$$

While voters are potentially more sophisticated, this assumption make estimation tractable. The model also assumes that $c'(g_j) = 1$, which one would expect in the absence of wasteful spending. This assumption further generalizes those in Calabrese et al. (2006).[24]

1.4.3 Equilibrium

An equilibrium consists of a set of communities, a continuum of households with a distribution of characteristics (α, y) each with housing and private consumption, a partition C_j across communities, and a vector of prices, taxes and public good expenditure such that:[25]

1. Every household maximize its utility subject to a budget constraint.
2. All households live in one community and do not want to move (sorting equilibrium).
3. Over all levels of feasible (g_j, t_j), at least half of the voters in j prefer the chosen levels over any other (voting equilibrium).
4. The housing market clears in all communities.
5. Each community's budget is balanced.

1.4.4 Implications of Spillovers and Centralization

The model provides a framework to understand the general equilibrium implications of spillovers and centralization. This section considers a two community example to illustrate the mechanisms. Consider communities 1 and 2 where $g_1 < g_2$, $p_1 < p_2$, and community 2

[24]Estimation of the structural parameters does not require assumptions on the cost function, only its derivative, $c'(g_j)$. However, when simulating alternative equilibria, changes in the cost for funding community college expenditure come from the reduced form results because they implicitly incorporate other changes in the equilibria like state appropriations and tuition payments. Other tax rates are fixed based on the results on K-12 education in Appendix A.3. These assumptions then do not require further specification of the functional form of the cost function or make additional assumptions on how state funding changes with alternative policies. One reason why K-12 spending do not seem to change could be due to Texas' elementary and secondary equalization formula, which is not the focus of this paper.

[25]Calabrese et al. (2006) prove that the equilibrium is unique conditional on observed community sizes and ranks.

is part of the taxing district while 1 is not. Community 1 therefore receives benefits that it does not help finance. Given the relative public goods and prices, household sorting is based on Equation 1.7. K summarizes the relative value of living in each area. A larger K implies that community 1 is more desirable and more households choose to live there.

$$\underbrace{\ln(Q_2 - Q_1)}_{\text{price difference}} - \underbrace{\ln(g_1^\rho - g_2^\rho)}_{\text{value difference}} = K$$

Positive public good spillovers lead g_1 to be relatively large. If they were removed the level of the public good in community 1 would decrease to $\hat{g}_1$. The sorting implications are summarized from how K changes with the level of the public good:

$$\frac{\partial K}{\partial g_1} = \frac{1}{Q_2 - Q_1}\left(\frac{\partial Q_2}{\partial p_2}\frac{\partial p_2}{\partial g_1} - \frac{\partial Q_1}{\partial p_1}\frac{\partial p_1}{\partial g_1}\right) - \frac{\rho g_1^{\rho-1}}{g_1^\rho - g_2^\rho} \tag{1.13}$$

The second right hand side term measures the change in the relative value of the public goods. Since ρ is negative and $g_1 < g_2$, this term is positive; a decrease in the public good in community 1 from the decrease or removal of spillovers will decrease K, all else equal. This term is the first order effect on the relative attractiveness of each community. At the initial prices, fewer households find it worthwhile to only receive $\hat{g}_1$ for p_1.

The first term on the right hand side captures the general equilibrium price response. When the public good in 1 decreases, more households move to community 2, which creates additional congestion and pushes up p_2. At the same time, congestion in 1 and p_1 fall. Taken together, the price response from the change in demand for each community is a second order effect that decreases the responsiveness of K.[26] Overall, $\partial K/\partial g_1$ is positive. While spillovers allow some households to free-ride, they also benefit the households providing the services by reducing equilibrium congestion and renters in both communities benefit from the spillovers.

The mechanisms for centralization are similar, except it additionally affects the tax rate.

[26]More formally, Q is increasing in p and therefore $Q_2 - Q_1 > 0$ since $\partial Q_j/\partial p_j = -Q\rho B p_j^\eta > 0$. $\partial p_1/\partial g_1 > 0$ and $\partial p_2/\partial g_1 \leq 0$ based on the first order sorting implications on congestion and the equilibrium conditions. Therefore the sign of this term is negative.

If community 1 joins the tax base, the level of the public good provided and tax rate should both weakly increase. A larger public good in 1 increases K but the increase in p_1 from the new tax acts in the opposite direction. The implications on households sorting and welfare are therefore ambiguous. Communities that receive very little or even no change in benefits from joining the tax base will be worse off. The change in the size of the tax base can also affect the public good level and the tax rate and the residents initially in 2. Annexations therefore redistribute across communities.

1.4.5 Brief Discussion of Model Limitations

The previous parts of this section present a tractable but highly stylized model of household sorting and public good provision to understand the general equilibrium welfare implications of spillovers and centralization. It is important to consider some of the key simplifications.

First, households in the model care about a single public good based on their preferences α. This implies that all households agree on the relative importance of each amenity and cannot sort based on how much they value a particular one. Even if benefits spill over to areas outside the tax base, it may not be welfare improving to force households who do not value the spending to finance it. In this case, the model would overestimate the value of spillovers and annexations. Since α varies across households, this feature is partially captured. Some have very little value of all public good spending, including by the college. However, as discussed in Section 1.2, the community college provides a broad public good by strengthening other services like K-12 education and community infrastructure.

The assumption on preferences and sorting implications of the model imply that communities are vertically differentiated, which limits the number and type of substitutes a household has. This may imply that the general equilibrium benefits or cost of annexing a particular community are more concentrated than in reality. Vertical differentiation also does not capture potential idiosyncratic preferences for communities and the model overstates mobility. I conduct all analyses in both partial equilibrium where households are immobile and general equilibrium where they are perfectly mobile to understand the implications.

The model's static framework of a single metro area also ignores some longer run benefits of increased access to community college services. If more households attend college because of centralization, this would lead to economic growth. Higher rates of attendance may therefore increase incomes, while lowering inequality. Overall, these effects are likely small, especially since take-up is costly for communities far from the campus, but would lead to an underestimate of the value of centralization. Additional economic growth may also attract more households to move to the metro area. In-migration would expand the tax base and potentially allow for more public good spending, but also lead to additional congestion.

1.5 Model Estimation and Identification

The structure of the model provides a set of estimating equations to recover the underlying parameters in three steps. First, I estimate the parameters governing household location decisions and housing expenditure, $\theta_1 = (\mu_{\ln y}, \sigma_{\ln y}, \lambda, \rho/\sigma_{\ln\alpha}, \eta, B)$. Conditional on θ_1, the public good moments identify $\theta_2 = (\rho, \sigma_{\ln\alpha}, \mu_{\ln\alpha})$. The housing supply elasticity, τ is taken from Saiz (2010). Finally, the structural parameters and policy variation pin down the relationship between the model implied public goods and amenities observed in the data.

The remainder of the paper focuses on the Austin area. In line with the assumptions of Tiebout (1956), where households do not face a work-residency restriction, it is natural to take a metropolitan statistical area (MSA) or commuting zone (CZ) as the household sorting area. For Austin, the MSA and CZ consist of the same set of counties. All 39 school districts that touch the area are included. This includes the entire Austin Community College (ACC) service area as well as part of the service areas of five other community colleges; Alamo, Blinn, Central Texas, Temple, and Victoria. Since school districts are important for household location choice and are the most relevant geography for taxing district annexation, the model treats them as communities.[27]

For each school district in the Austin area, the 2015 American Community Survey (ACS)

[27] Service areas and taxing districts are primarily made up of school districts, although this is not always the case. Cities or counties can be annexed as well, and these boundaries do not necessarily line up with the school districts.

reports the number of households, the marginal distribution of household income, and home and rental values.[28] From the ACS data, I convert home values to rental values following Poterba (1992). I additionally estimate the price of a unit of housing services for each ISD in a hedonic regression following Sieg et al. (2002, 2004) using parcel level sale and house characteristic information from CoreLogic Tax Roll Record File for 2015.

I combine the cross-section of data on household sorting with a panel of community amenities and spending to capture the effects of changes over time. Integrated Postsecondary Education Data System (IPEDS) reports yearly community college instructional expenditure per student, the Common Core of Data reports the yearly ISD instructional expenditure per student, and the the Uniform Crime Reporting Statistics and the Texas Department of Public Safety report the yearly crime rate. See Appendix A.2 for more information about the data.

The remaining parts of this section, I show how the model implies a set of equations that identify the structural parameters, θ. Quantities labeled with θ note that they are computed from the model rather than observed in the data.

1.5.1 Household Sorting

The model implies both the population shares, and the income distribution of households within each community. The first set of moments match ISD income quantiles conditional on correctly predicting n_j. From Equation 1.8, the size of community 1 is

$$\underbrace{n_1}_{\text{data}} = \int_{-\infty}^{\infty} \int_{-\infty}^{K_1 + \rho \frac{y^{1-\nu}-1}{1-\nu}} f(\ln \alpha, \ln y) d \ln \alpha d \ln y$$

where $K_0 = -\infty$. Households with both the lowest incomes and lowest preferences for the public good will live in the community with the lowest price of housing and lowest level of public good provision. Since the ACS reports n_1, I can back out K_1 conditional on θ_1. After estimating K_1, I apply the same procedure to find the remaining K_j conditional on K_{j-1},

[28] 1-year estimates from the ACS are used when available. For smaller school districts, the 5-year estimates are used. Data is from IPUMS (Manson et al. 2019)

and θ_1. These community specific intercepts are equilibrium parameters that are functions of prices and public goods. The procedure to find L_j From Equation 1.10 is analogous.

The previous expression implies that the qth quantile of the income distribution in community j is given by

$$\underbrace{q}_{\text{picked}}\underbrace{n_1}_{\text{data}} = \int_{-\infty}^{\ln y_j(q,\theta)} \int_{-\infty}^{\hat{K}_1+\rho\frac{y^{1-\nu}-1}{1-\nu}} f(\ln\alpha, \ln y) d\ln\alpha d\ln y$$

where $\hat{K}$ denotes that it is estimated based on the observed population shares. Since the ACS reports the 20th, 50th, and 80th percentiles of the income distribution, the estimation procedure matches those model implied log income quantiles to the data to form three estimating equations, one for each quantile q:

$$\ln y_j(q,\theta) = \ln y_j(q) + \epsilon_j^y(q) \tag{1.14}$$

where ϵ_j^y is mean zero measurement error. These equations identify $\mu_{\ln y}, \sigma_{\ln y}, \lambda, \rho/\sigma_{\ln\alpha}$ and ν.

1.5.2 Housing Market

Taking logs of Equation 1.4 and noting that the rental value of housing, r, equals the price times quantity of housing demanded, gives

$$\ln r_{jq} = \ln(B_t) + (\eta+1)\ln p_j + \nu\ln y_j(q,\theta) + \epsilon_{jq}^r \tag{1.15}$$

where ϵ_{jq}^r is mean zero measurement error. Using the same three quantiles as before, this equation creates an additional set of moments that identify η and B, as well as ν.

1.5.3 Public Good Provision

The model provides two predictions about the level of the public good, given the observed household location decisions and prices. First, rewriting Equation 1.7, the level of the public good in each community $j > 1$ can be estimated as follows

$$g_j^s = \left\{ g_1^\rho - \sum_{i=2}^{j} (Q_i - Q_{i-1}) \exp(-K_{i-1})) \right\}^{1/\rho} \tag{1.16}$$

where g_j^s is the model implied public good based on the implications from the sorting equilibrium. The locus of pivotal voters from Equation 1.11 implies that

$$g_j^v = \left\{ Q_j B_t p_j^\eta \frac{dp_j}{dg_j} \exp(-L_j) \right\}^{1/(\rho-1)} \tag{1.17}$$

I estimate the mean and standard deviation of the preference distribution to minimize the difference between these two predictions:

$$g_j^s = g_j^v + \epsilon_j^g \tag{1.18}$$

where $\mathbb{E}[\epsilon_j^g | p_j, H_j] = 0$. Changes in the preference distribution will change the set of indifferent households between two adjacent communities, I_j, and the locus of median voters and therefore differentially affect the implied sorting equilibrium and median voters' optimal policies. This strategy finds the preference distribution such that the two model implications are closest across communities.[29]

1.5.4 Estimation and Results

From the previous section, there are seven model implied equations; three income quantiles, three housing expenditure quantiles, and one public good. The structural parameters are identified using cross-sectional variation from household decisions. The three income quantile and three housing expenditure equations estimate θ_1 using weighted non-linear least squares. I then estimate $\theta_2 | \theta_1$ from the public good equation. Table 1.4 displays the estimates.

Based on these results, the single-cross property holds since ρ is negative. The remaining

[29] The procedure is similar to Calabrese et al. (2006), but does not require additional assumption on the cost of providing the public good because I estimate home prices, p_j, from sales data.

estimates are similar to the previous literature using this approach; for example, the correlation between income and preferences for the public good is small and negative. Lower income households may have stronger preferences because they are less able to purchase substitutes, for example private K-12 education. η, the housing demand elasticity, is slightly larger in magnitude than the estimate of -0.116 found for Los Angeles in Sieg et al. (2004), which also estimates the price of housing services from home sale data.[30] The estimate of the housing supply elasticity, τ, is taken from Saiz (2010), who finds that it is 3 for the Austin MSA.

1.5.5 Public Good Decomposition

The model's structure and the estimated parameters inform how households respond to changes in public goods and taxes across space and recover the equilibrium level of the public good under the decentralized status quo policy. To measure how the public good would change under different policies, I assume the composite public good, g, production function is Cobb-Douglas in community amenities observed to the researcher, X, including community college spending per student, ISD spending per student, and the crime rate:

$$g_{jt} = \left(\prod_k (X_{jt}^k)^{\pi_k} \right) \exp(u_{jt})$$

where g_{jt} is an equilibrium parameter from the model and X is data. u_{jt} is a community-year specific amenity that is additively separable in a community amenity, ϕ_j, a time amenity ξ_t, and a community-year term ϵ_{jt}^g. Therefore

$$\log(g_{jt}) = \pi_{1jt} \log(\mathrm{CC}_{jt}) + \pi_2 \log(\mathrm{ISD}_{jt}) + \pi_3 \log(\mathrm{Crime}_{jt}) + \phi_j + \xi_t + \epsilon_{jt}^g \tag{1.19}$$

where π_1 is a function of distance to the campus and whether a community is in the taxing district: $\pi_{1jt} = \pi_1^1 + \pi_1^2 \mathbb{1}(\text{Not in TD}) + \pi_1^3 \text{distance}_j + \pi_1^4 \text{distance}_j \times \mathbb{1}(\text{Not in TD})$. Since π_1

[30] Other estimates of η using this methodology without home sale data range are relatively large in magnitude. Walsh (2007) estimates η is -0.62 and Epple and Sieg (1999) estimate it is -0.7.

is approximated to be linear in distance, I also assume that it is non-negative.[31] Equation 1.19 estimates the contribution of community college spending across the service area and taxing district to the public good using observed policy variation from the annexations and changes in community college spending over time, conditional on other amenities that affect household location decisions. By using two-way fixed effects, the πs are estimated using within community variation, but they can also be interpreted as amenities themselves. For example, if households value living near a river, as in Austin ISD, then this feature will be captured by ϕ. ξ captures the value of amenities in a given point in time, like the quality of public transportation in a given year or state level policies.[32]

For each public good component, X, the production function determines how changes in that component affect the aggregate public good, or $\partial \log g / \partial \log X$. The elasticity can be computed from

$$\underbrace{\left(\frac{\partial \log p}{\partial \log X}\right)}_{\text{amenity capitalization}} = \overbrace{\left(\frac{\partial \log p}{\partial \log g}\right)}^{\text{public good capitalization}} \underbrace{\left(\frac{\partial \log g}{\partial \log X}\right)}_{\text{amenity contribution}} \tag{1.20}$$

since prices in the model are a function of the public good, which is a function of amenities through the production function.

The first term on the right side, $\partial \log p / \partial \log g$, measures the elasticity of the price of a unit of housing with respect to the public good. This term captures how changes in the public good are capitalized into housing prices. When the public good increases in j, the community becomes more attractive and additional households want to move in, which

[31] Allowing the parameter to be a higher order function of distance does not affect the implications much. Also note that $\phi_j = \sum_i \pi_5^i \mathbb{1}(i = j)$, and ξ_t can be defined similarly for year fixed effects.

[32] The functional form assumptions provide a way to focus on the key features for understanding the implications of community college expenditure policy. However, they likely obscure from other potentially relevant features. For example, community college expenditure may produce spillovers across service areas. In reality, these are likely small since those distances are large. The Cobb-Douglas assumptions also restrict the relationship between amenities (Albouy et al. 2020). A high quality community college may a strategic substitute or complement for ISD spending, but as a first-order approximation, these interaction effects are ruled out.

pushes up the price of housing. The extent of which depends on the structural parameters, especially the elasticity of housing supply, τ. I estimate this term by simulating the general equilibrium change in price in each community j for a 1 percent increase in g_j financed by additional property taxes, holding g_{-j} fixed. The elasticity in this setting is about 0.6; a 1 percent increase in the public good leads to a 0.6 percent increase in the price.

$\partial \log p / \partial \log X$ reflects how changes in amenities affect housing prices, as in the event-study in Section 1.3. I estimate the elasticity for each public good component from Equation 1.19 with the log of the Census Tract repeat sales home price indices as the dependent variable, using all of the policy variation from across the state not just the Austin area. ϕ_j are estimated with Census Tract fixed effects.[33] Estimating this specification at the Census Tract level provides a fine level of geography to understand the heterogeneous treatment effects of spending and annexation by distance to the campus.

This strategy builds on the previous literature to credibly estimate the value of amenities using causal inference methods. Previous work using this type of model, following Epple and Sieg (1999), rely on functions of the income rank of communities as instruments in a cross-section to jointly estimate these parameters and the preference distribution. The instrument is valid in the model since housing prices and public goods monotonically increase in income rank. Instead, the approach used here estimates $\partial \log p / \partial \log X$ using policy variation in panel data, and then calculates the model implied outcome– how much must have the public good increased to be consistent with the observed change in home prices in the data? The estimates are reported in column (1) of Table 1.5 along with the scaled public good production function parameters in column (2).

The production function estimates imply that increases in community college quality improve the public good, but that this benefit is concentrated near the campus. The value of spending decays slightly more quickly for those not in the taxing district and therefore

[33]The repeat sales home price captures of the value of living in a particular Census Tract by removing home characteristics, like the number of bedrooms. See Appendix A.2 for more information about the home price and amenity data.

the value of annexation is increasing with distance, up to a point; the public good benefit from joining is an inverted-U function of distance.[34] Communities more than 38 miles from the campus do not benefit from the college, based on this specification, while the average ISD that is currently not part of a taxing district is 42 miles away (Table 1.1). Expanding the tax base by annexing communities far from the campus puts a tax burden on some areas without an increase in benefits. In 2016, Brazos County residents attempted to start their own community college, rather than join their service area's taxing district, citing distance as a primary reason. However, in the ACC Service Area, the boundary of the taxing district is only 17 miles away. Marginal changes in the taxing district boundary would reduce free-riding. This suggests that ACC's tax base may be too small, although the estimates here do not capture the general equilibrium implications of household and housing market responses.[35]

1.6 Measuring the Value of Spillovers

The relative value of centralized and decentralized policy setting importantly depends on the geographic nature of the benefits. The decomposition estimates imply that spillovers are relatively large on the current boundary of the taxing district but decrease to zero. I quantify the value of spending and spillovers by simulating household compensating variation (CV) for a 10 percent increase in ACC's expenditure (about $610 per student) financed by an increase in property taxes.[36] After this policy change, households in j receive $(\hat{g}_j, \hat{p}_j)$. Letting V_0 denote a household's indirect utility in the baseline, and V_1 in the counterfactual

[34] The inverted-U shape is determined by the signs and magnitudes of π_1^3 and π_1^4. Since π_1^3 is negative, the benefits of being in the taxing district decrease with distance from the college to 0. It is large enough in magnitude that benefits reach 0 within the service area. A negative and small π_1^4 implies that benefits decrease slightly more quickly outside of the taxing district and so the benefits from joining are increasing with distance as long as $\pi_1^1 + \pi_1^2 \text{distance}_j$, or the value of being in the tax base, is positive.

[35] Note that, in general, there is a negative relationship between the crime rate and the price index, but this almost entirely captured by the Census Tract and year fixed effects.

[36] The increase in taxes is calculated from summing over the communities' budget constraints in the taxing district using the facts that the marginal cost of the public good is 1 and the tax rate is uniform.

with higher spending, partial equilibrium CV is then computed as:

$$V_0(g_j, p_j, y, \alpha) = V_1(\hat{g}_j, \hat{p}_j, y + \text{CV}, \alpha) \tag{1.21}$$

CV in this setting is the amount of additional income (in 2015 dollars) a household would need to receive to be just as well off after the policy change. A positive value implies a household is worse off after a 10 percent increase in spending and associated increase in taxes. This partial equilibrium measure reflects the short-run effects before the households can adjust their location decisions. The implied welfare change per capita in j, W_j, is

$$W_j = -\left[\int_{C_j} \text{CV}\ f(\ln\alpha, \ln y) d\ln\alpha d\ln y / n_j + R_j^0 - R_j^1\right] \tag{1.22}$$

where R_j^s is the per capita housing rents to landowners in state of the world $s \in \{0, 1\}$, $\int_0^{p_j^h(s)} H_j^s dp/n_j$. Welfare is reported so that a positive number implies a community is better off after the policy change.[37]

Panel (a) of Figure 1.2 presents the partial equilibrium per capita welfare change for a 10 percent increase in ACC expenditure in each ISD based on households' baseline residency. A darker shade denotes that a community benefits more from the increase in spending. The map is divided into three parts; communities labeled "TD" are part of ACC's tax base, those labeled "N" are in ACC's service area but not in the taxing district, and those without a label are not part of the ACC service area but are in the Austin MSA.[38] In partial equilibrium, only communities in the ACC service area have a change in their public goods or taxes, and therefore a welfare change.

All but one community in the taxing district is worse off, on average. The increase in the public good is not worth the additional tax burden to finance it. If the college taxing district optimally chooses the level of the public good to maximize its residents' welfare,

[37] Using an equivalent variation measure instead of CV leads to quantitatively very similar estimates.
[38] Appendix A.5 decomposes the welfare change into the effect on rents, housing prices, and public goods.

we should expect this result. One community in the taxing district, however, is better off because it is only 5.5 miles from the campus. It is also the wealthiest and residents have the strongest preferences for the public good. The variation in the welfare change across the taxing district highlights an important cost of centralization– not all communities agree on the optimal level of spending because of community variation in household characteristics (Tiebout 1956; Alesina et al. 2004). With spatially delineated public goods and a uniform tax rate, the distributional effects are especially important because increases in spending redistribute from less wealthy communities far from the campus to the wealthy nearby.

The increase in expenditure benefits those outside of the taxing district. Communities near the campus have especially large per capita gains of up to about \$775; they receive more public goods at no additional cost. However, the benefits decrease with distance from the campus and the four farthest ISDs in the service area receive no additional benefits. There is substantial amount of free-riding on the border of the taxing district, but none far away.

Since household locations are fixed in partial equilibrium, the previous results do not reflect how spillovers affect the housing market. Increases in spending and the tax rate increase the benefits from free-riding and, in general equilibrium, households move in response. Similar to Equation 1.21, the general equilibrium CV is

$$V_0(g_j, p_j, y, \alpha) = V_1(\hat{g}_{j'}, \hat{p}_{j'}, y + CV, \alpha) \tag{1.23}$$

for a household who lives in j in the baseline and j' when spending increases.

In general equilibrium, public goods and taxes both adjust from the changes in expenditure and household location decisions. If the increase in spending and taxes leads some households to leave the taxing district, the size of the tax base will shrink. The college must raise the rate, based on the elasticity from Table 1.3, to keep the level of spending constant. The higher rate puts more of a burden on the tax base and pushes additional households to leave, compounding the effect, while congestion works in the opposite direction. In principle, changes in the size of the tax base from sorting could also affect expenditure, but the estimates from Table 1.3 suggest they do not.[39] Panel (b) of Figure 1.2 shows that the per capita general equilibrium welfare gain for each ISD based on the household's 2015 residency.

When households are mobile, they leave areas with a new large tax burden relative to the change in the public good. This has two effects on welfare for the tax base. First, some households move to a community that better suits their preferences, which drives down CV and increases W_j. Second, the households that remain benefit from lower housing prices and less congestion, although the tax rate slightly increases because the tax base shrinks by about 1 percent. This benefits renters at the expense of landowners.

Of the 17 communities in the ACC service area but not in the taxing district, a 10 percent increase in expenditure increases welfare in all but one. Additional expenditure particularly benefits the communities outside the taxing district but close to the campus. On average, these households value the increases in public goods more than the negative effects on pricing from congestion. CV is therefore positive but lower in magnitude than the partial equilibrium case. The higher housing prices additionally benefit landowners. Communities farther from campus have lower gains from additional spending but still face higher prices. The partial equilibrium estimates therefore understate the value of spillovers on communities far from the campus because they do not capture the housing market externalities.

The increase in ACC's expenditure also imposes a \$27 welfare loss per capita for those outside the service area due to the general equilibrium externalities. These areas become relatively less attractive; many areas now have higher public good levels with no new tax burden. As households leave, landowner profits decrease. Some low-income households are also made worse off because the higher spending in ACC reduces the number of low public good, low price communities. Increased competition for the remaining ones drives up prices.

[39] We may expect spending per student to increase because of decreases in average fixed costs or administrative costs. After an annexation, average fixed costs may not decrease because the college expands the geographic scope of its activity. Administrative costs likely do not decrease because increases in the tax base and annexations do not change the cost of assessing homes or collecting taxes. After a Danish reform that increased jurisdiction size, Blom-Hansen et al. (2016) also find no increase in economies of scale.

Overall, the welfare effects in general equilibrium are slightly smaller in magnitude than in partial equilibrium but these two approaches attribute the burden to different groups. In the short run, the policy redistributes from communities in the taxing district relatively far from the campus to the closest community to the campus and those just outside the boundary of the tax base. When households are mobile, sorting pushes some of the losses onto landowners and more onto low-income communities in the tax base.

1.7 Centralized Community College Provision

Who benefits, and by how much, from centralization? I now quantify the value of centralized provision for ACC using the compensating variation method described by Equations 1.21 and 1.23. Currently, 9 ISDs are part of ACC's taxing district but 26 are in the service area. Under centralization, the taxing district includes all 26, which is more than a 23 percent increase in aggregate property value. With the larger tax base, the college lowers the statutory rate based on the reduced form elasticity from Table 1.3 and communities joining the taxing district receive weakly higher public goods based on the decomposition estimates. The partial equilibrium per capita welfare gains for centralization are presented in Panel (a) of Figure 1.3. Households in the baseline 2015 taxing district labeled "TD" have very modest welfare gains of \$8.50 on average from the slight decrease in taxes. Since higher income households consume more housing, the benefits are concentrated in wealthy areas.

The welfare gains for newly annexed communities are an inverted-U function of distance. Communities close to the campus receive very large spillovers and do not benefit much from joining the taxing district but face a new tax burden. Before annexation, the college was easily accessible. Additional resources spent in the district after annexation are substitutes, rather than new benefits worth the cost. As distance from the campus increases slightly, so do the benefits from joining. Communities between 21 and 38 miles from the campus had less access to the college's services before annexation and now benefit in excess of the new tax burden from the additional resources. The four communities more than 38 miles from the campus have welfare losses from centralization of between \$93 and \$177. They have an additional tax burden without a compensating increase in public goods.[40]

In partial equilibrium, the ISDs outside the ACC service area are unaffected. When households can sort after the expansion of the tax base, the changes in household location decisions will affect equilibrium housing prices and the distribution of welfare gains.[41] To account for these externalities, the general equilibrium welfare gains from centralization are shown in Panel (b) of Figure 1.3 and decomposed in Panel (b) of Figure 1.4.

Centralization leads to a general equilibrium yearly welfare gain of \$23M for the MSA or \$29 per household but those benefits are not equally spatially distributed. Households that originally live in the newly annexed communities have an average benefit of about \$160. The increase in the level of the public good makes their communities more desirable, which increases housing prices and therefore rents to landowners but makes renters worse off from the added congestion. Comparing Panels (a) and (b) of Figure 1.4 for the annexed areas, the general equilibrium effects on both housing prices and public goods are smaller in magnitude than when households are immobile. Some households are priced out of their original communities and move to lower (g, p) neighborhoods. Annexation has the opposite effect on the areas very close to or very far from the college. With almost no increase in the public good but a new tax burden, the ISD becomes less desirable. Congestion and rents therefore decrease, which also mitigates the welfare losses from tax increases.

Households in the baseline taxing district also have a small welfare gain from centralization valued at \$8 per capita. When the newly annexed become more attractive, congestion in these communities decreases but this has about equal in magnitude welfare consequences on prices and rents. Although these mechanisms sum to about zero for W_j, they imply redistribution from landowners to renters. Some also benefit from moving to the newly annexed areas that now have better public goods, although this effect is small.

[40] Panel (a) of Figure 1.4 breaks down the welfare changes by community and source of benefit type.

[41] Changes in household location decisions and prices jointly impact the size of the aggregate community college tax bases and therefore also affect equilibrium tax rates. In practice, this effect is small.

Centralization additionally imposes negative externalities on the communities outside ACC's service area. These areas become less attractive and housing rents decrease. The public good and price welfare effects are driven by household sorting. Some households originally outside ACC service area now find it worthwhile to move into it where they pay higher housing prices for more public goods.[42] On average, the welfare effects from these two sources are roughly equal in magnitude, but the price effect is especially large for households in low median-income communities. After centralization, these areas become more vertically differentiated. The change in the choice set for low-income and low-preference households leads many to live in worse matches compared to the baseline because low-price-low-public good communities become a scare resource.[43]

Given the potential benefits from annexation, why is the tax base relatively small compared to the total size of the service area? The results from this section suggest that the contiguity requirement plays an important role. Communities close to the campus have small gains or even loses from annexation since the benefits are an inverted-U shape in distance. If only free-riding communities on the boundary are eligible to join, then it is perhaps not surprising that tax bases do not expand.

However, there are some communities on the boundary with positive welfare gains from annexation. They may not join because of information frictions or perceptions about the benefits. The new tax is salient, but they may not know how joining the tax base will improve the public good in their communities. Households may also believe that the new revenue will mainly benefit those close to the campus, in contrast with the results found here, which would imply that annexation is a form of tax exporting. Incorrect beliefs likely play an important role since community colleges spend resources before the annexation to

[42]The households that remain outside of the ACC service area have also experience a housing price change from resorting, but the sign of the changes varies across the distribution of communities.

[43]In the baseline, the lowest housing price and public good area has a very large amount of land and is in ACC's service area. Since it is 31 miles away from the campus, its public goods and therefore housing price increase a lot after annexation. Many households were originally able to live in a very low (g, p) community, but now loose this option. Couture et al. (2019) also find that endogenous amenities from rising income inequality leads to similar welfare effects on the low-income.

educate the voters about the public good benefits, although that is outside the scope of this analysis. Households far from the campus may also be reluctant to join if they believe spending or the tax rate will increase in the future, for example because of declining state funding, as highlighted in Section 1.6.

1.8 Welfare Implications of the Property Tax Base Size

Under the current configuration of ACC's taxing district, some households free-ride. This leads to low public good provision in many communities, but also decreases congestion in the tax base. Centralization may not be optimal since expansions redistribute from the wealthy communities close to the campus to the less wealthy far away[44] or because it generates negative externalities from household sorting.

How much should the tax base "optimally" expand? Let S be the set of possible boundary configurations, then the optimal tax base size is:

$$S^* = \underset{s \in S}{\arg\max} \int \mathcal{W}(V(\alpha, y, g(s), p(s)), R(s), j(s)) f(\ln \alpha, \ln y) d \ln \alpha d \ln y \tag{1.24}$$

for social welfare function $\mathcal{W}$, which depends on households' indirect utility and their locations, and landowner rents. For example, the social planner may incorporate the welfare of all households in the MSA and set $\mathcal{W}(\cdot, j) = \mathcal{W}(\cdot)$, whereas ACC may only care about households in its service area and ignore the externalities on others.[45]

To measure the welfare implications of changes in the property tax base size, I simulate the welfare effects of different boundary configurations for a utilitarian planner. Since the tax base must be contiguous by state law, I divide the service area into three parts; the 2015 taxing district, and inner ring of communities currently in the service area but not the taxing district, and an outer ring. The inner ring consists of 11 communities that are currently

[44] In 2015, the correlation between distance and median community income is -0.45.

[45] This method of computing social welfare has a closed form solution which makes it much more tractable than using compensating variation as in the previous sections. However, it requires that housing rents are attributed to households. I assume housing rents in j are equally shared so that each household who lives in j in the baseline receives the same amount.

eligible to join the taxing district because they border it while the outer ring consists of 6 that border the inner ring but not the taxing district. Within each ring, I calculate the new equilibrium and total social welfare for every possible boundary configuration. There are currently 9 communities in the taxing district, which implies that there are 11 possible ways to have a 10 community tax base using only the inner ring, and so on.[46]

In partial equilibrium, the social planner considers the net benefits to each community from annexation and the benefits to the tax base from a lower statutory rate. Households do not move in response to the policy and there are no additional externalities created. Allowing some communities close to the campus to free-ride could increase welfare if the cost of the new tax burden after annexation is more than the benefits of the slight reduction of the tax rate. Similarly, annexing far away communities that receive no benefits could increase welfare if the reverse is true. Panels (a) and (b) of Figure 1.5 plot the percent welfare gain from the different boundary configurations considered. The social planner values additional annexations because many households benefit from the added spending in their area more than the new tax burden and expansions decrease the burden on the current taxing district.

Not all increases in the size of the tax base increase aggregate welfare. Annexing communities near the campus only slightly increases their public goods and are not worth the new tax burden, even after accounting for the decrease in the tax rate on the other households. Since the college is easily accessible before joining the tax base, annexation can lead to wasteful spending rather than an improvement in public good quality.

In general equilibrium when households move in response to the new distribution of public goods and tax rates, a change in the taxing district's boundaries can create positive or negative externalities on other communities and community college service areas nearby. The welfare effect of the externality depends on features of the communities being annexed, like the amount of land, and of their closest substitutes. Panels (c) and (d) of Figure 1.5 plot the welfare gains from different configurations allowing households to move and housing

[46] The number of configurations is 2^{11} for the inner ring and 2^{6} for the outer ring, giving 2112 total.

prices to adjust. The welfare effects are larger in magnitude and more varied than in partial equilibrium since the sign of the externalities is ambiguous. In 113 cases, the negative externalities are so large that welfare decreases from the status quo decentralized policy.

In general equilibrium, centralization is not socially optimal. A tax base that is one community smaller raises per capita welfare by about $222 even though the residents in the removed community benefit since it is about 35 miles from the campus. This highlights that the externalities created from household sorting in response to changes in spatial public good provision are first order. The removed community is just above the median of the price and public good distribution. After annexation, its public good increases a lot and it acts as another good substitute for high-price-high-public good locations, which pushes prices in those communities down. High income communities are more harmed by the loss in landowner profits than households who remain benefit from paying lower housing prices from reduced congestion.

The change in sorting and congestion creates negative welfare effects on lower income communities. Some households are crowded out of the middle-ranked communities and into the lower public good and price areas. The increase in aggregate demand for low ranked communities increases housing prices, where the lower income face a larger burden from the change in the cost of living than they benefit from high rents to landowners. More affordable, although low public good, communities become a scarce resource. Finally, the removed community has little land and few residents so its gains do not affect aggregate welfare as much as the externalities imposed on the remaining metro area. Of the 2112 boundary combinations considered, 146 increase social welfare over centralization, with tax bases as small as 12 communities.

1.8.1 Returns to Scale and Property Tax Base Size

When the tax base expands, the college has additional revenue that it can use to lower the tax rate or increase spending. The amount of additional funds depends on the extent to which there are increasing returns to scale. The estimates from Table 1.3 imply that that there is only a small decrease in the tax rate, but no change in expenditure per student after an increase in aggregate property value. A larger elasticity has an ambiguous effect on the relative value of centralization. If additional spending mainly benefits high-income areas, it can reduce the negative sorting externalities since fewer households would find it worthwhile to move to the middle-income areas. Instead, if the spending mainly benefits the middle-income neighborhoods, the externalities will be exacerbated. In that case, removing communities from the tax base could lead to higher spending levels.

The implications of increasing returns to scale are shown in Figure 1.6. It plots the percent welfare gain over decentralization for both centralized provision and the highest welfare 25 ISD boundary configuration from Figure 1.5 under different assumptions on the elasticity of expenditure per student. As the elasticity increases, so does welfare in both cases since the level of the public good provided in many communities rises without an added tax burden. At the same time, the welfare gap grows– increasing returns to scale compound the negative externality effects and undermine the value of centralization.

1.9 Conclusion

Public good provision by different levels of government has important implications for equity and efficiency, especially in the presence of spillovers. This paper considers the relative merits of different tax base configurations to fund community colleges. Evidence from recent expansions suggests that all communities benefited from a higher level of government providing the services since aggregate home values and prices rose. However, the costs and benefits are not uniformly distributed across space. Even in partial equilibrium when there are no externalities from household sorting, not all potential annexations are welfare enhancing. The state's uniform tax requirement puts too much of a burden on distant communities that receive very little in benefits even after they join the tax base and creates other wasteful spending in nearby areas. In general equilibrium, changes in the geographic scope of the net benefits in Austin Community College's service area create externalities that undermine the value of a large geographic tax base. Centralization is therefore not socially optimal.

The results highlights two important general equilibrium considerations for policymaking. First, changes in the level of government provision can create both positive and negative externalities from household sorting that alter the distribution of benefits and burdens to fund the public good. While centralization may help solve the free-riding problem, it has other consequences for redistribution. In particular, the expansion of Austin Community College's taxing district leads to large negative welfare effects on low-income areas outside of the college's service area. Second, under current regulations, the college and voters likely do not account for the externalities when deciding whether to expand. They may not care about the costs imposed on other households. The state can play an important role as an even higher level of government to help correct them.

The policy analysis considered in this paper follows current Texas law. The tax rate must be uniform, the tax base must be contiguous, and the state is unable to collect property taxes, limiting the scope of centralization. Since the tax rate to fund community colleges is only about five percent of the total property tax bill, the uniform rate requirement plays a small role for optimal policy. However, because the benefits from annexation are an inverted-U shape in distance, the contiguity requirement implies that only those with relatively low gains from joining can. In 2016, Michigan passed no longer required community college tax base to be contiguous, although this has not lead to any more successful annexations.

Future work should study other public goods with different benefit and spillover production processes. In this setting, the public good is produced in a single location by a coalition of communities that choose one level of expenditure. Other spatial amenities with externalities, like pollution and crime, involve more heterogeneity across communities in production, which has an ambiguous effect on the returns to increasing the property tax base size. Centralization would limit the ability of households with different preferences to sort, but would also limit the production of negative externalities.

Table 1.1: Summary of ISDs by Date Joined Taxing District

	Has Not Joined	After 1991	Before 1991
Panel A: 2016 School District Statistics			
Population	11,780	78,784	61,298
Percent with Some College	16.0	12.9	15.8
Percent with Associates' Degree	4.5	3.7	5.3
Percent with College Degree	12.8	13.6	14.3
Median Income ($)	54,160	55,139	60,390
Median Home Value ($)	118,883	138,021	143,339
Median Rent ($)	572	680	715
Aggregate Home Value ($ Millions)	535	3,839	2,943
Percent Own Residence	59.1	56.7	56.3
Number of Businesses	288	1,088	962
Distance to Main Campus (Miles)	41.9	19.7	12.2
Panel B: 2016 Aggregate Statistics			
Percent of ISDs	69.4	4.9	25.8
Percent of Land	83.3	3.2	13.5
Percent of Population	29.4	13.8	56.8

Note: Panel A presents summary statistics for Texas school districts based on the year they join a taxing district. The first column presents the means for school districts not currently in a taxing district, the second for those who joined after 1991 and the final for those who joined before 1991. 1991 is chosen because it is the start of the Texas Comptroller data. Panel B presents aggregate statistics to describe overall coverage of the taxing districts.

Table 1.2: The Effect of Annexations on other Communities

	Initial Tax Base		Remaining Service Area	
	(1) Log Home Value	(2) Log HPI	(3) Log Home Value	(4) Log HPI
Log # Other School Districts in TD	0.273	0.079	0.335***	0.066**
	(0.179)	(0.055)	(0.084)	(0.022)
School District FE	Yes	No	Yes	No
Census Tract FE	No	Yes	No	Yes
Observations	442	9522	2158	6795

Note: Columns (1) and (2) show the estimated effect of the number of school districts in the community college taxing district on log of aggregate single family home values and home price indicies (HPI) respectively for the initial tax base, that is communities that were part of the tax base in 1991 when the sample begins. Columns (3) and (4) estimate the effects on the remaining communities in the service area who are not part of the taxing district by 2016. All specifications include year fixed effects. Standard errors are cluster bootstrapped with 999 replications, clustered at the school district level. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.3: The Effect of Tax Base Expansions on Community Colleges

	(1) Log College Tax Rate	(2) Log Instructional Per Student
Log Aggregate Market Value	-0.204***	0.013
	(0.060)	(0.031)
Observations	1031	1019

Note: Column (1) shows the estimated effect of the elasticity of the community college statutory tax rate with respect to changes in the aggregate market value of property in the tax base. Column (2) shows the estimated effect of the elasticity of instructional expenditure per student. All specifications include year and community college fixed effects. Standard errors are clustered by community college. The sample includes 50 colleges for 21 years (1998-2018). Differences in sample sizes are due to missing data. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.4: Structural Parameters

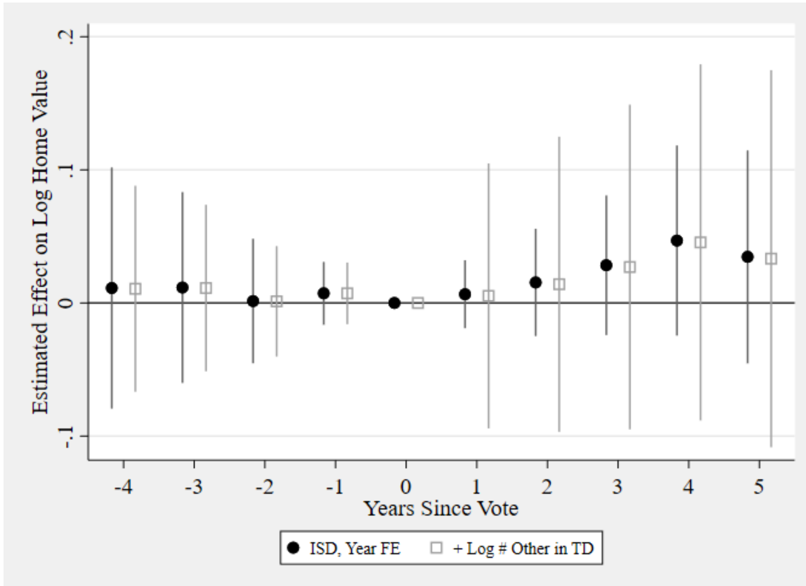
Name	Parameters	Estimate	SE
Stage 1:			
Mean log income	$\mu_{\ln y}$	10.897	0.005
Sd log income	$\sigma_{\ln y}$	0.794	0.006
Income-preferences correlation	λ	-0.066	0.005
Income elasticity	ν	0.976	0.002
Elasticity of substitution/Sd log preferences	$\rho/\sigma_{\ln \alpha}$	-0.316	0.007
Price elasticity	η	-0.226	0.119
Housing Demand	$\ln B$	-1.174	0.062
Stage 2:			
Mean log preferences	$\mu_{\ln \alpha}$	-1.956	0.191
Sd log preferences	$\sigma_{\ln \alpha}$	0.290	0.073
Elasticity of Housing Supply	τ	3	-

Note: The structural parameters are estimated in two steps. Heteroskedastic-robust standard errors correct for the two-step procedure.

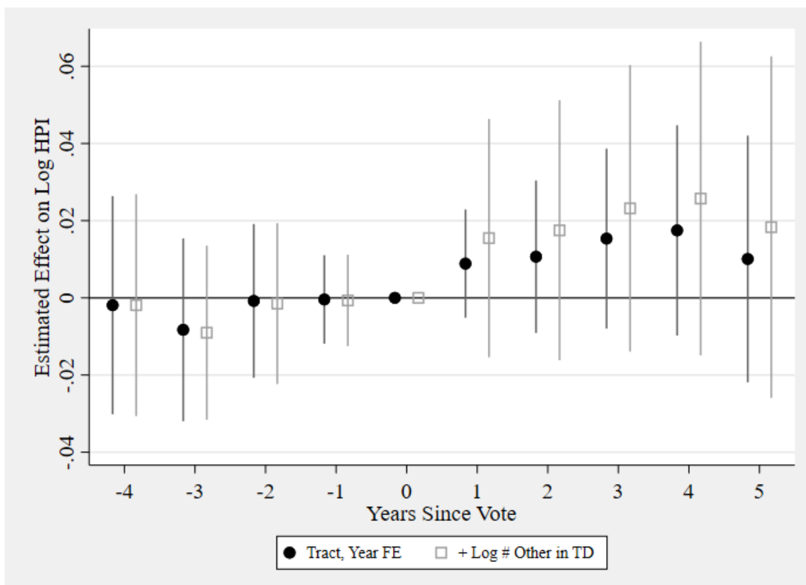
Table 1.5: Public Good Production Function Decomposition

	(1) $\log p_{jt}$	(2) $\log g_{jt}$
Log Expenditure/Student	0.422***	0.720***
	(0.114)	(0.194)
Log Expenditure/Student $\times$ $\mathbb{1}$(Not in TD)	0.004	0.007
	(0.003)	(0.006)
Log Expenditure/Student $\times$ Distance	-0.011**	-0.019**
	(0.004)	(0.007)
Log Expenditure/Student $\times$ Distance $\times$ $\mathbb{1}$(Not in TD)	-0.0003^{+}	-0.0005^{+}
	(0.0001)	(0.0002)
Log ISD Spending/Student	0.169***	0.289***
	(0.061)	(0.087)
Log Crime Rate	0.006	0.011
	(0.023)	(0.040)
Observations	6808	6808

Note: Column (1) estimates the effect of log amenities on the log of Census Tract home price indices. The specification includes year and Census Tract fixed effects. When data are missing, for example, because school district spending data was not reported, we include an indicator variable to maintain the full sample. Standard errors are computed by cluster bootstrap with 999 replications, clustered by ISD. Column (2) scales the results from Column (1) by 1/0.6 to compute the public good elasticities. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$



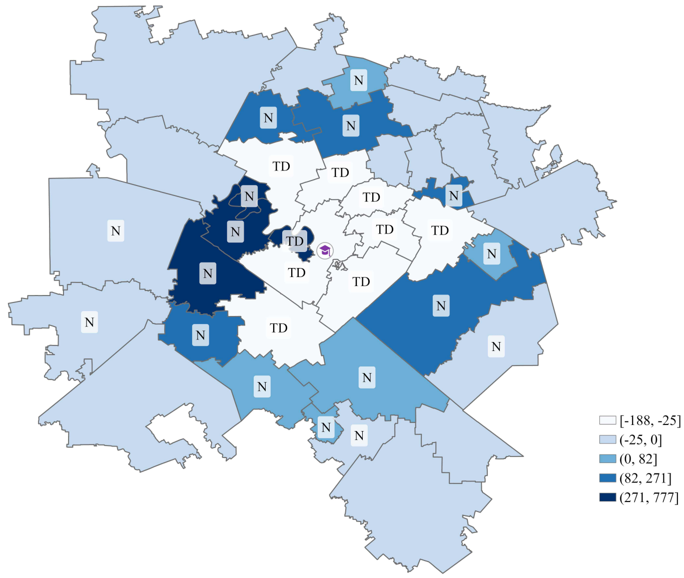
(a) Effect on Home Values



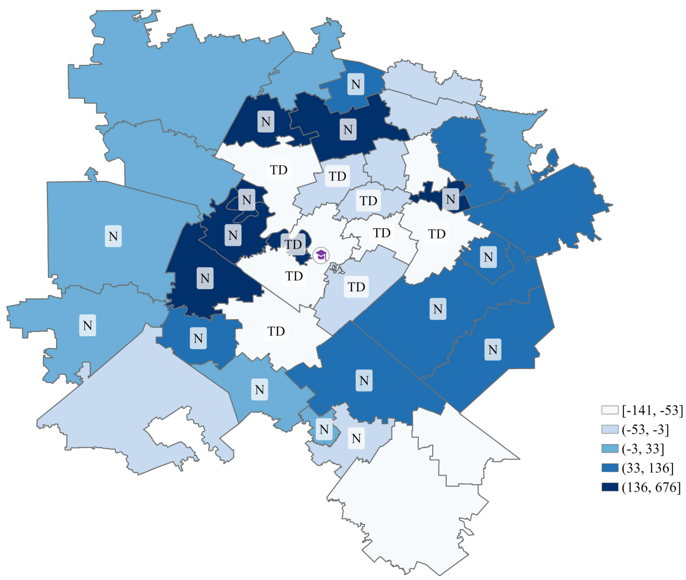
(b) Effect on HPI

Figure 1.1: Effect of Joining the Taxing District on Housing Values and Prices

Note: Panel (a) plots the event-study coefficients and 95 percent confidence intervals for the effect of joining the tax base on the log of aggregate single family home values in the school district using the sample of school districts that are annexed between 1995 and 2011. The sample size is 1114. The solid black circles estimate the effect including year and school district fixed effects, and the clear squares additionally control for the log of the number of other school districts in the tax base. Panel (b) plots the coefficients and confidence intervals for the effect on the log of HPI in the Census Tract including year and Census Tract fixed effects. The sample size is 6808. Standard errors are cluster bootstrapped with 999 replications, clustered at the school district level in all specifications.



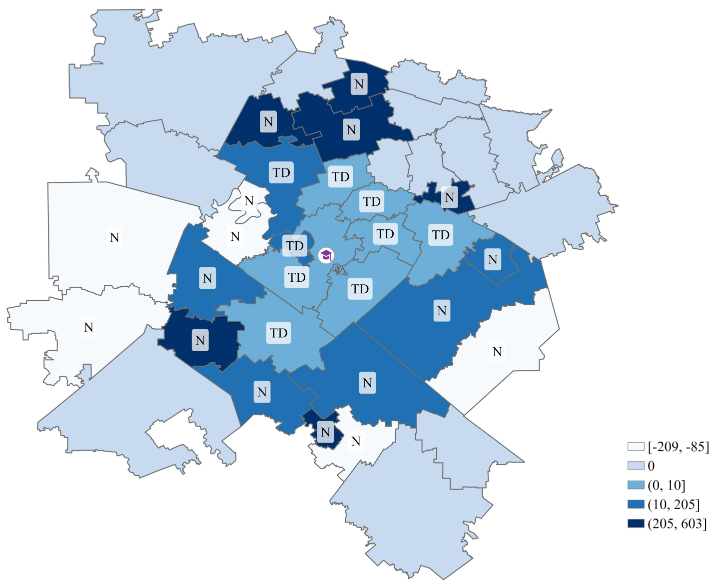
(a) Partial Equilibrium



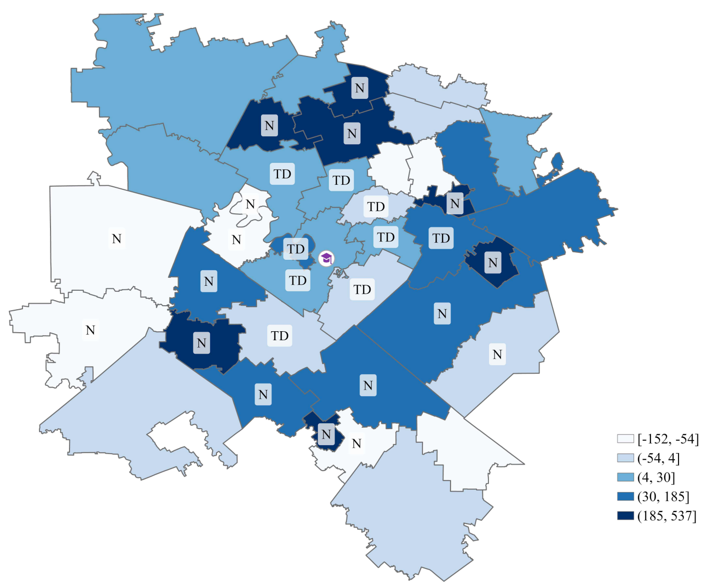
(b) General Equilibrum

Figure 1.2: Per Capita Welfare Gain for a 10 Percent Expenditure Increase

Note: Panel (a) shows the partial equilibrium per capita welfare gain for a 10 percent increase in instructional expenditure per student where household are immobile and (b) shows the general equilibrium results where households are perfectly mobile. Each community is shaded according to the simulated per capita welfare gain of residents by their baseline location. Darker shades indicate larger welfare gains. The location of Austin Community College is marked with a graduation cap. Communities labeled TD are in Austin Community College's taxing district, N are in the service area but not taxing district, and the unlabeled are not in either but in the MSA.



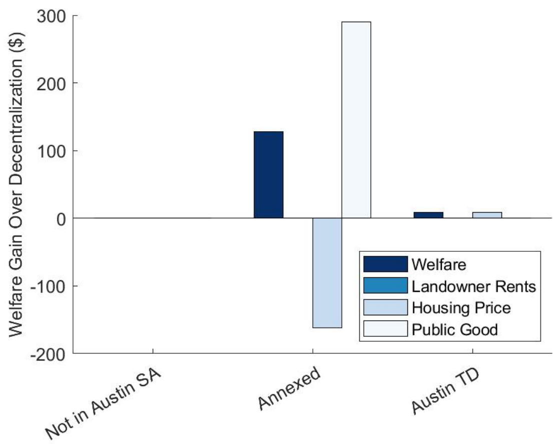
(a) Partial Equilibrium



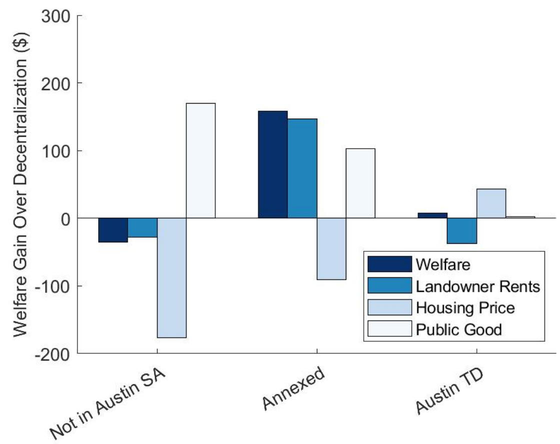
(b) General Equilibrum

Figure 1.3: Welfare Gain for Centralization

Note: Panel (a) shows the partial equilibrium per capita welfare gain from centralization where all 26 school districts in Austin Community College service area are part of the taxing district fixing households' original location and (b) shows the general equilibrium results when households are perfectly mobile. Each community is shaded according to the simulated per capita welfare gain of residents by their baseline location. Darker shades indicate larger welfare gains. The location of Austin Community College is marked with a graduation cap. Communities labeled TD are in Austin's taxing district, N are in the service area but not taxing district, and the unlabeled are not in either but in the MSA.



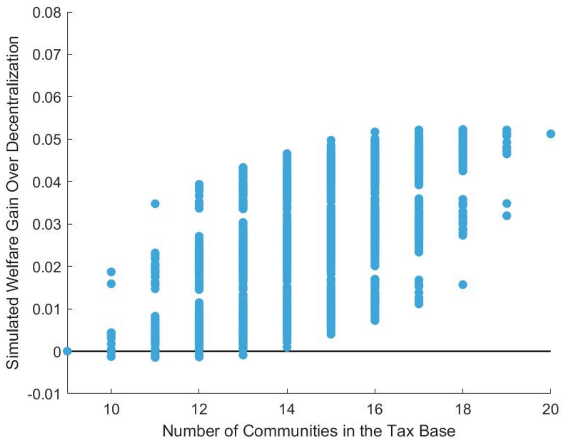
(a) Partial Equilibrium



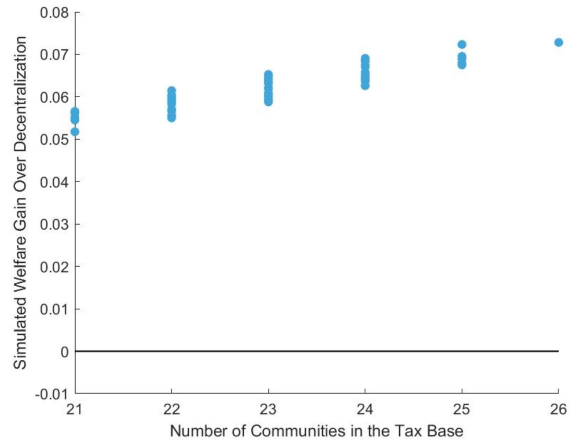
(b) General Equilibrium

Figure 1.4: Welfare Decomposition

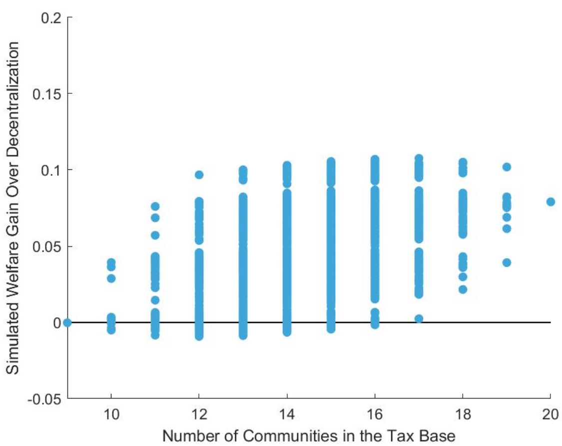
Note: Each figure is divided into the three community types: those in the current Austin Community College taxing district, those annexed through centralization, and those not in the Austin Community College service area but in the Austin MSA. Panel (a) decomposes the partial equilibrium welfare change for centralization when households are immobile and Panel (b) shows the general equilibrium change when they are perfectly mobile. The left most bar plots the average welfare per capita of baseline residents. The next three bars break down welfare into its components, housing rents to landowners, compensating variation for the housing price change only holding public goods fixed, and compensating variation for the public good change holding the price of housing fixed.



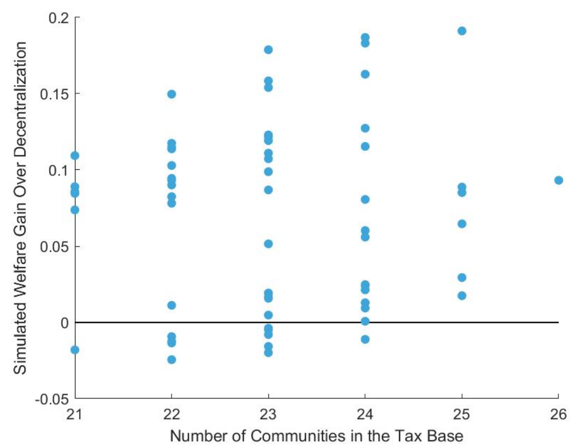
(a) Inner Ring Partial Equilibrium



(b) Outer Ring Partial Equilibrium



(c) Inner Ring General Equilibrium



(d) Outer Ring General Equilibrium

Figure 1.5: Welfare Implications of the Property Tax Base Size

Note: Panel (a) shows the welfare gain for each possible boundary configuration of the taxing district using the 11 communities that border the baseline tax base (inner ring) holding residencies fixed in partial equilibrium where households' residencies are fixed. This gives $2^{11} = 2048$ configurations. Panel (b) conducts the same exercise for the 6 communities that do not border the baseline tax base (outer ring), where all communities in the inner ring also join. This gives $2^6 = 64$ configurations. Panels (c) and (d) compute the general equilibrium gain when households are perfectly mobile. Each circle represents the percent welfare increase over the baseline for all households.

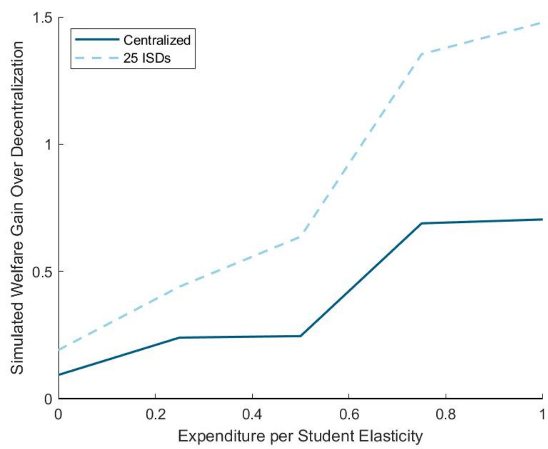


Figure 1.6: Returns to Scale, Welfare, and the Property Tax Base Size

Note: The figure plots the general equilibrium model implied percent welfare gain over the 2015 Austin Community College tax base as a function of the instructional expenditure per student elasticity with respect to aggregate property value. The solid dark line shows welfare under centralization with all 26 school districts in the taxing district, while the dashed light line shows the welfare with 25 school districts.

Chapter 2

Optimal Minimum Wage Setting in a Federal System (with Matthew Wilson)

2.1 Introduction

After Kansas City, and St. Louis, Missouri set their own minimum wages above the state level in 2017, the Missouri state legislature prohibited any city from setting its own policy. While there is likely a political motive for a red state to overturn a progressive policy in its blue cities, there are also potential economic benefits from centralized redistribution. The central government can more efficiently implement such policies when workers are mobile by internalizing spillovers, and a state may not want one of its cities to set a higher minimum wage because of the externality imposed on the rest of the state. This may justify minimum wage preemption laws in Missouri and 24 other states. However, since the state is constrained to setting a uniform policy, as in Oates' (1972) decentralization theorem, then local policy setting may be preferred.

We examine optimal minimum wage setting in a federation with mobile workers to understand the relative benefits of centralized, decentralized, and joint policy setting with interregional spillovers. Which level of government should set minimum wage policies? Although our substantive focus is on the minimum wage, our framework has implications for a broad set of policies where local governments supplement federal decisions, like the EITC, income taxes, and Medicaid. Understanding the appropriate level of government decision making is increasingly important as the US federal government has continued to shift more responsibility for these policies to states (Baicker et al. 2012).

Our work most directly builds on Lee and Saez (2012), who examine the role of the minimum wage in the competitive labor market of a single jurisdiction, where workers are implicitly immobile. They find that a binding policy is desirable, even when non-linear taxes are available, if the newly unemployed have zero surplus from working and the government values redistribution to low-wage workers. In their model, workers can migrate between low and high wage jobs, though they face different costs from working in each. When a worker's sector is fixed so that labor supply can only respond on the extensive margin, the minimum wage is second-best Pareto inefficient.[1] With only one jurisdiction, their framework is unable to consider the relative merits of having different levels of government set policy as well as the externalities that result from horizontal and vertical government competition. Vertical competition in a federation is particularly important for the minimum wage because it is a price floor policy; only the higher one matters. If the central government that is restricted to a uniform policy sets a binding minimum wage in the lower wage region, then it lowers the costs associated with horizontal competition.

We adapt the Lee and Saez (2012) model to a two-jurisdiction framework with mobile agents, regional governments, and a federal government to study the trade-off presented in the decentralization theorem. Workers are not mobile across sectors, but the high-skilled are mobile across jurisdictions.[2] Our goal is to understand the conditions under which setting a binding minimum wage is optimal for each type of government and analyze the welfare implications of different policy setting authorities. Local governments compete for high-skilled workers, which hinders their ability to redistribute through the minimum wage, but the central government is restricted to a uniform policy. Our stylized framework only includes two inputs for tractability, but the high-skilled can be thought of more broadly as

[1] Gerritsen and Jacobs (2016) allow for educational investment in the Lee and Saez framework and find that a minimum wage is only optimal in the presence of non-linear taxes if the gains from more education outweigh higher unemployment.

[2] Blundell and MaCurdy (1999), and Lee and Saez (2012) both note that the extensive margin working decision where workers are not mobile across sectors is the most important and relevant in practice.

mobile factors of production that may respond to increases in the minimum wage, like capital and firm location. In addition to the "tiered" US model of a federal uniform price floor in which states may "top off," or raise the floor, we consider federal uniform, decentralized, and federal non-uniform policy setting.

We first provide theoretical results to understand the extensive margin decision of implementing a binding minimum wage when governments maximize social welfare. The local governments will set a binding policy if the welfare loss from unemployment is zero on the margin, and the government values redistribution to low-skilled workers more than emigration. Our conditions are the same as in Lee and Saez, except governments are concerned with policy induced migration. The federal government only cares about migration insofar as it affects total output and moving costs, leading it to have stronger preferences for a minimum wage than local jurisdictions.[3] Since the theoretical model only gives predictions on the extensive margin, we proceed to calibrate a two region model to the aggregate US economy to illustrate how mobility and regional heterogeneity in productivity impact the relative benefits of centralization, decentralization, a hybrid system as in the US, and centralized non-uniform policy in general equilibrium.

The key insight from the model is that different levels of government are strategic complements in policymaking and the extent of this complementarity depends on input mobility and regional heterogeneity in productivity. Higher levels of mobility increase the costs associated with horizontal competition, while higher levels of regional heterogeneity decrease the effectiveness of uniform central policy. When jurisdictions are identical and inputs are mobile, centralized authority leads to greater social welfare while additionally allowing local governments to enact policy has no effect. However, when jurisdictions differ but inputs are immobile and therefore there are no interregional spillovers, decentralized authority is preferred. Simultaneously allowing the central government to also set a minimum wage in

[3]We also discuss joint policy setting by the local and federal government as well as federal non-uniform policies and several alternative models, such as rent-seeking governments, economy-wide aggregate production functions, and additional mobile factors of production, to understand how the model assumptions impact the sufficient conditions for a minimum wage.

this case does not change the equilibrium. For the more realistic cases with heterogeneous regions and imperfectly mobile inputs, a hybrid system improves welfare since central uniform policy reduces horizontal competition and decentralized policy allows for a different minimum wage in each jurisdiction.

Regional heterogeneity in redistributive preferences may also drive differences in the minimum wage, as reflected in the Missouri example. Urban areas with high minimum wages tend to be both more productive and more progressive. We find that progressivity has a nonlinear effect on optimal policy since the minimum wage redistributes to the low-skilled from the high-skilled and the newly unemployed. A government with a very progressive social welfare function would not implement a binding minimum wage because of the negative effects of additional unemployment, while a government at the other extreme would not implement a binding policy because it does not value redistribution. Local heterogeneity in preferences can also lead to higher minimum wages. If one local government enacts a minimum wage because of its redistributive preferences, this can lead the other local government to set a binding policy, since a higher minimum wage in one region decreases the migration externality in the other.

To understand the welfare implications of tiered and centralized policy setting, we calibrate an economic geography model of the continental United States to match regional heterogeneity in productivity, high-skilled location decisions, employment, and the federal government's optimal policy. Consistent with our earlier findings, this model predicts that tiered policy setting yields a small welfare gain over centralization alone, as it allows the federal government to more effectively redistribute from high-skilled to low-skilled workers. Even though states are heterogeneous in productivity, and high-skilled workers are fairly mobile, we also find that tiered minimum wage setting closely approximates the social planner's optimal policies. Similar models have been used to study place based policies in general (Kline and Moretti 2014), corporate tax cuts (Suárez Serrato and Zidar 2016), and income taxes (Colas and Hutchinson 2020).

While we focus on the trade-offs of different minimum wage setting authorities, the previous normative literature focused on different tax systems. Stiglitz (1982), Allen (1987), and Guesnerie and Roberts (1987) find that the minimum wage is not optimal when nonlinear taxes are available, however, Guesnerie and Roberts find they may be desirable under linear taxes. The literature has also considered more sophisticated minimum wage policies in a single jurisdiction where government competition is implicitly absent, like a graduated minimum wage tied to firm size (Danziger and Danziger, 2018), in-kind redistribution (Economedies and Moutos, 2017), monitoring job search (Broadway and Cuff, 2001), and bargaining power (Hungerbuhler and Lehmann, 2009).[4] Previous work on optimal tax that incorporates government competition in a federation (e.g. Wilson, 1982; Gordon, 1983; Hamilton and Pestieau, 2005; Gordon and Cullen, 2012; Divorkin, 2017) has not simultaneously allowed governments implement minimum wages. We connect these two previous literatures by understanding optimal minimum wage policies as a tool for redistribution in a federation.[5]

By putting the optimal minimum wage analysis into a federation, our work relates to the literature on tax competition. In our model, when a jurisdiction increases its minimum wage, high-skilled workers migrate to the other region, creating a horizontal externality that is increasing in mobility. Similar to models like Wilson (1986), and Zodrow and Mieszkowski (1986) on tax competition, agent mobility will lead to low local minimum wages. By adding a federal government to the Wilson-Zodrow-Mieszkowski set-up, Keen and Kotsogiannis (2002, 2004) find vertical externalities leave state taxes too high, which is in contrast with our result that federal policy decreases the cost of local minimum wage setting. Our results are therefore more similar to Janeba and Wilson (2011) who examine local public good provision with both horizontal and vertical externalities. In their model, local provision is

[4] We examine the implications in a competitive labor market, but other work has focused on the optimality when there are search frictions, as in search-and-matching model. See Flinn (2006), Dube et al. (2016)), and Lavecchia (2018), for example.

[5] There is also a large empirical literature on the effects of the minimum wage. See Neumark (2018) for a recent review.

too low due to competition, but central provision is inefficient because it is determined by a winning coalition in the legislature. In other political economy models, like Lockwood (2002) and Besley and Coate (2003), central provision is inefficient for similar reasons, whereas it is inefficient in our framework because the federal government is restricted to a uniform policy. Our assumption is based on the US system, but this restriction does not need to hold. In May 2019, European Commission Vice-President Frans Timmermans called for each member of the European Union to set a minimum wage of 60% of its median salary.

Although not centered on the minimum wage, there is a complementary literature on the interaction of mobility and redistribution, beginning with Stigler (1957) and Oates (1972). Our focus in this work is on the implications of different levels of government undertaking redistribution as in Oates (1977), Ladd and Doolittle (1982), Brown and Oates (1987), and Dixit and Londregan (1998). Our approach to understanding this mechanism is similar to Epple and Romer (1991), who also use a computational model to capture general equilibrium effects. The importance of mobility for policy making in a federation has also been studied in the context of public goods (Epple and Platt 1998; Caplan et al. 2000; Calabrese et al. 2012; Simon 2020),[6] income redistribution (Wildasin 1997; Armenter and Ortega 2011; Gordon and Cullen 2012), and higher education (Wildasin 2000).[7]

2.2 Model

We consider a simple two-region, two-factor model to emphasize the role of regional heterogeneity and factor mobility in determining the relative effectiveness of centralized and decentralized minimum wage setting. The economy, I, consists of two regions, $i \in \{1, 2\}$, low-skilled and high-skilled workers, local firms, and local and federal governments. We consider each in turn.

Each region i has a unit mass of immobile low-skilled labor. Low skilled labor may

[6] There is also a relatively large literature on the value of centralized environmental policies with spillovers. See Banzhaf and Chupp (2012) for a recent example.

[7] There is also a growing empirical literature on income taxation and mobility. See Kleven et al. (2019) for a review.

either be employed, l, or unemployed, u, so that $u_i + l_i = 1$. If employed, they pay a cost θ of working, which is distributed according to a known distribution $\mu(\theta)$. There are $\bar{h}$ high-skilled workers who are (im)perfectly mobile across regions, so that $h_1 + h_2 = \bar{h}$. High-skilled workers face moving costs of $\xi \sim \zeta(\xi)$, but do not face a working cost.[8] We interpret this as a normalization on working costs so that the low-skilled costs are relative to those of the high-skilled. It also ensures that all high-skilled workers are in the labor force for any reasonable minimum wage policy. Workers can only be employed in their region of residence. Consumption, c_j, for each worker j is given by her wage minus her cost of working and moving costs paid, if applicable.

Our mobility assumptions are based on the high correlation between education level and mobility, and the fact that low-skilled workers sluggishly respond to employment shocks.[9] To better understand the implications of these restrictions on workers, we consider how low-skilled mobility affects our theoretical results in the next section. Then, in section 2.5, our calibration exercise varies $\zeta(\xi)$ to describe the relationship between the high-skilled migration elasticity and optimal minimum wage setting authority. Although the model and calibration are based solely on low and high-skilled workers for tractability, the high-skilled represent input factor mobility more broadly. For example, firms may respond to a minimum wage in their capital investment or location decisions and local governments compete for these resources. Our choice of using the high-skilled as the mobile factor allows us to naturally incorporate the mobile input into a social welfare calculation. We discuss how additional mobile inputs affect our results later in this section.

Each region has a single perfectly competitive firm with production function $\gamma_i f(l, h)$. Without loss of generality, we assume $\gamma_1 = 1$ and $\gamma_2 = \gamma \leq 1$. Workers are paid their marginal product, and if the high-skilled are perfectly mobile, they migrate so that $w_1^h = w_2^h$. Relative

[8] In our simplified framework, we abstract away from congestion costs, which are likely important for high-skilled migration (Moretti 2013).

[9] Amior (2019) documents this relationship using the CPS. Our assumptions are similar to those in Wildasin (2000) who examines the public provision of human capital in a federation, but we allow for variable high-skilled mobility to understand the comparative statics. We discuss this assumption more explicitly later in this section.

low-skilled competitive wages across regions depends on γ and the relative stock of high-skilled workers. We assume that the optimal minimum wage is never binding for high-skilled workers, so that they are also always employed. Let Θ_{ie} be the set of θ's of the low-skilled employed workers in i and Θ_{iu} the unemployed. For a given wage, workers with relatively large values of θ are unemployed because they choose to exit the labor force. If there is a binding minimum wage, others become unemployed due to the change in low-skilled labor demand by the firm. Similarly, define Θ_{ih} as the set of high-skilled workers in i.

There is one central government, and each region has its own local government. Both the local and central governments have perfect information about workers' working and moving costs and regional productivity, and we explore the consequences of this assumption in Section 2.5.3. Governments care about the consumption of their constituents, according to the concave function $G(c)$, where $G' > 0, G'' < 0$, and $G'(0) < \infty$. Total utility in r is given by $\int_{j \in r} G(c_j)dj$, $r = \{1, 2, I\}$, i.e. each region maximizes the total utility of its own residents, while the central government maximizes the utility of both regions. We assume that $G(c_h) > 0$ for any feasible policy so that high-skilled workers are valued and local governments compete for them. Federal and local governments are either endowed or not with the ability to set a minimum wage. There are four potential authority structures: (1) only regional governments; (2) only central government uniform; (3) both central uniform and regional governments; and (4) central non-uniform. Governments do not need to raise revenue.

Equilibrium is determined in two stages. In the first stage, no government has minimum wage setting authority; this is the competitive equilibrium. This gives each worker a residence. In the second stage, some governments are endowed with minimum wage setting power and optimally set their policy, given the policies of other governments, if applicable, and the responses by firms and workers.

Definition 1. *The economy is in equilibrium after the second stage for a given authority structure if:*

1. *Each government with minimum wage setting authority optimally sets its minimum wage to maximize the welfare of its residents, given the policies of the other governments with authority.*

2. *Firms maximize profits, taking the residencies of workers and government policies as given.*

3. *Each worker optimally chooses whether to enter the labor force to maximize utility.*

4. *No high skilled workers want to move.*

2.2.1 Decentralized Minimum Wage

We consider the conditions under which it is optimal for local governments to set a binding minimum wage. Each jurisdiction i's social welfare function is defined as

$$SW_i = (1 - l_i)G(c_u) + \int_{\Theta_e} \mu(\theta)G(c_l)d\theta + \int_{\xi(\Theta_{ih})} \zeta(\xi)G(c_{h_i})d\xi \tag{2.1}$$

where $c_j = w_j - \theta_j$ for low-skilled employed individual of type θ_j, $c_h = w_h$ for high-skilled individuals, and c_u is the consumption of the unemployed. Without taxes and transfers, $c_u = 0$. Since moving costs are only paid if residents leave, they do not directly enter the regional government's social welfare function. As in Lee and Saez (2012), let $g_u = G'(c_0)/\lambda; g_l = \int_{\Theta_e} \mu(\theta)G'(c_l)d\theta/\lambda$ and $g_h = G'(c_h)/\lambda$, where λ is the Lagrange multiplier on the Social Planner's budget constraint.[10]

For a given level of the minimum wage in the other jurisdiction, $\bar{w}_{-i}$, i implements a binding minimum wage $\bar{w}_i$ if and only if:

$$0 < \left.\frac{\partial SW_i}{\partial \bar{w}_i}\right|_{\bar{w}_{-i}} \iff \tag{2.2}$$

[10]The Social Planner's budget constraint is $(1 - l_i)c_u + l_i c_l + h_i c_h \leq l_i w_l + h_i w_h$. Since the governments in our model do not redistribute and only set a minimum wage, they do not have a budget constraint. Also note that the construction of the g terms are analogous for the regional and central governments, although their respective populations are different.

$$
\begin{aligned}
0 < & \frac{-\partial l_i}{\partial \bar{w}_i} G(0) + \frac{\partial l_i}{\partial \bar{w}_i} \int_{\Delta_{\bar{w}_i}\Theta_{iu}} \mu(\theta) G(c_l) d\theta \\
& + \int_{\Theta_{ie}(\bar{w}_i)} \mu(\theta) G'(c_l) d\theta \frac{\partial w_i^l}{\partial \bar{w}_i} + h_i G'(c_h) \frac{\partial w_i^h}{\partial \bar{w}_i} \\
& + \frac{\partial h_i}{\partial \bar{w}_i} G(c_h)
\end{aligned}
$$

The first line represents the change in welfare induced by newly unemployed low-skilled workers. $\Delta_{\bar{w}_i}\Theta_{iu}$ denotes the set of workers who become unemployed due to the minimum wage $\bar{w}_i$, which depends on how workers are separated from the firm after the policy change. Following Lee and Saez (2012), we assume efficient rationing, i.e. workers with the lowest surplus from working involuntarily lose their jobs first. Under the assumption, $\Delta_{\bar{w}_i}\Theta_{iu} = [\gamma_i f_l(l,h), \theta(\bar{w})]$, where $\theta(\bar{w})$ is the largest θ worker who becomes unemployed. $\theta(\bar{w})$ depends on the new policy level, the relative production functions in each jurisdiction, $-i$'s minimum wage, and the amount of migration. For a very small binding policy, $\theta(\bar{w}) = \gamma_i f_l(l,h)$. Therefore, the marginal low-skilled worker who loses her job has $c_j = 0$ before and after the policy change. The second line represents the change in welfare from the still-employed in region i and the third line is the change in welfare due to emigration.

First note that $\frac{\partial w_i^l}{\partial \bar{w}_i} > 0$ by construction and $\frac{\partial l_i}{\partial \bar{w}_i} < 0$ since it is less profitable to employ low-skilled workers at a higher wage. Increasing the minimum wage also impacts the high-skilled workers since $\frac{\partial w_i^h}{\partial \bar{w}_i} < 0$ from the zero-profit condition.[11] As the marginal product of labor for high-skilled workers falls, some high-skilled workers leave, giving $\frac{\partial h_i}{\partial \bar{w}_i} < 0$ as well. For this reason, the third line of the first order condition only involves emigration and not immigration. Since the local government only cares about its residents, i.e. natives less those who move out, it does not internalize individual migration costs.

Proposition 1. *If (1) there is efficient rationing; (2) the demand elasticity is finite; (3) the supply elasticity of low-skilled workers is positive; (4) the government values additional*

[11] With perfectly competitive markets and constant returns to scale, the zero-profit condition implies that $ldw^l + hdw^h = 0$, giving the relationship between the minimum wage and high skilled wages. See Lee and Saez (2012) Appendix A.1. for more information.

wages to low-skilled workers more than the loss of wages to high-skilled (redistribution) and outmigration, then it is total welfare improving for a jurisdiction to impose a (small) binding minimum wage.

Proof. Our assumptions imply that the first two terms sum to 0 and the magnitude of the third term is larger than the sum of the magnitude of the fourth and fifth terms. More explicitly, using the above expressions and the zero profit condition, we can rewrite the last three terms as $l_i\lambda[g_l - g_h] + \frac{\partial h_i}{\partial \bar{w}_i} G(c_h)$. This is identical to proposition 1 in Lee and Saez (2012) except jurisdictions are now also affected by potential outmigration, weakening the incentive to enact a binding minimum wage. □

Under these assumptions, this extensive margin decision to adopt a minimum wage depends on γ, $\bar{w}_{-i}$, the elasticity of substitution between low and high-skilled workers, and the mobility of the high-skilled. In section 2.4.2, we conduct a calibration exercise to better understand how these parameters impact equilibrium, on both the extensive and intensive margins.

We now consider how changes to the assumptions of Proposition 1 affect the desirability of a binding minimum wage. Without efficient rationing, the first two terms do not sum to zero for a small increase in the minimum wage because workers with a large surplus from working lose their jobs. If workers were fired randomly after the increase in the minimum wage, then social welfare would be lower in expectation as workers with positive value of employment are separated rather than the indifferent marginal employee. However, even if the lowest cost workers were separated first, a government may still implement a binding minimum wage if the value of redistribution is sufficiently large.[12] This reasoning leads to a corollary to Proposition 1.

[12] Lee and Saez (2012) note that while this is the most favorable assumption for optimal policy, it may not be realistic or may be costly to reach through queuing or search costs. However, there is some empirical literature supporting this assumption. Neumark and Wascher (2007) find that the minimum wage has larger impacts on teenagers and secondary earners who are more likely to have more elastic labor supply and Luttmer (2007) found that reservations wages do not increase with the minimum wage.

Corollary 1. *Efficient rationing leads local governments to weakly prefer the highest minimum wages, all else equal, relative to any other separation assumption.*

Proof. Efficient rationing maximizes the sum of the first two terms, while not affecting any other terms. Otherwise, there is a first order welfare loss from unemployment. The condition is weak because a government may not desire a binding minimum wage for any separation policy if the other conditions of Proposition 1 do not hold. □

If conditions (2) or (3) do not hold, then the employment of low-skilled workers will change too dramatically for the first two terms to sum to zero. Condition (4) may fail if governments care too much about outmigration relative to redistribution, if they do not care about redistribution, or if they only care about redistribution to the unemployed. As a special case, consider when high and low-skilled labor are perfect substitutes. Then $\frac{\partial w_i^h}{\partial \bar{w}_i} = 0$ and there is no outmigration. In this case, the minimum wage is optimal if there is efficient rationing as the government trades off the unemployment of marginal workers with the additional earnings of all others. As the two types of labor become less substitutable, increasing the minimum wage will lead to lower high-skilled wages and more outmigration. Decreasing the elasticity of substitution will weaken the incentive to have a minimum wage, all else equal.

To further understand the migration externality, consider the impact of raising the minimum wage on the other jurisdiction. If high-skilled workers migrate from jurisdiction i because moving costs are sufficiently small on the margin, the marginal product of labor of high-skilled workers in jurisdiction $-i$ will fall. Government $-i$ is made relatively better off, all else equal, as there is a small change in high-skilled consumption, but a larger change in welfare due to the additional high-skilled workers. This effect lowers the cost for $-i$ to implement a binding minimum wage as fewer high-skilled workers will want to migrate compared to the case where i does not have a binding minimum wage. In our model, the marginal product of low-skilled workers increases when high-skilled workers immigrate, and wages and employment both increase, leading unambiguously to higher utility for those workers.

2.2.2 Mobility and the Minimum Wage

Throughout the model we assume that low-skilled workers are perfectly immobile for tractability. While these workers tend to have relatively low mobility rates overall, Monras (2019) finds that after a minimum wage increase, the relative share of low-skilled workers decreases because of changes in low-skilled in-migration. This effect is potentially driven by immigration from other countries. Cadena (2014) provides evidence that low-skilled immigrants prefer states with unchanged minimum wages relative to those experiencing increases because of the disemployment effects. Since our static framework considers a fixed point in time without population change, including from immigration, or shocks to workers that induce migration besides directly from the minimum wage, this concern is not particularly relevant in our highly stylized setting. Monras additionally finds no effect on out-migration. Taken together, these effects on migration are consistent with moving low-skilled not choosing to live in minimum wage areas, rather than migration due to changes in the minimum wage itself. In a dynamic setting where households may move at any given time for any reason, the immobility assumption becomes less plausible and more important.

We now consider the implications of relaxing the perfect immobility assumption.[13] Low-skilled workers initially sort in stage 1 so that w^l is equal across jurisdictions. After a minimum wage increase in i, the firm does not hire any additional low-skilled and some are separated. Therefore, no low-skilled in $-i$ have an incentive to immigrate to i after the policy change. The unemployed in i emigrate if the new wage is higher than the sum of working and moving costs. After a minimum wage increase, low-skilled employment weakly increases in $-i$ from high-skilled migration, but the model does not specify who is hired. If the firm in $-i$ only hires initially local low-skilled workers, then there is no incentive to move and the immobility assumption does not impact our results.[14] If the firm hires workers with the

[13]Fukumura and Yamagishi (2020) study government competition with low-skilled migration responses from minimum wage changes. However, they focus on an economy with only minimum-wage workers and identical regions.

[14]Phillips (2018) finds that low wage employers discriminate against applicants with a long commute, which provides suggestive evidence for this assumption.

lowest working cost, then the incentive to move depends on γ. Absent any binding minimum wage policies, the marginal worker in jurisdiction A will have a weakly higher working cost than the marginal worker in B since $\gamma \leq 1$. Therefore, a minimum wage in A will not lead to low-skilled migration, but one in B will if moving costs are sufficiently small. In general, adding in low-skilled mobility has an ambiguous effect on the optimal minimum wage policy that depends on the value of $G(0)$ and whether the newly unemployed would be hired if they moved.[15]

The main migration mechanism in the model is the link between regions created from high-skilled. The extent to which high-skilled workers respond to local minimum wage policies, $\frac{\partial h_i}{\partial \bar{w}_i}$, determines the effectiveness of local government policy relative to centralization. If the derivative is zero, because moving costs are high or the marginal product of low and high-skilled workers are unrelated, then our framework reduces to the single region case as studied by Lee and Saez (2012). Monras (2019) estimates the relationship between wages and high-skilled out-migration, finding negative or zero effects; however, they are not statistically different from zero. Cadena (2014) also provides some suggestive, although mixed, evidence that this derivative is in fact negative. Taking the point estimates across all specifications at face value, he finds a negative relationship between the count of immigrants with at least some college and the minimum wage in a given state.[16] More generally, the high-skilled migration externality represents the impact on endogenous business locations of low-skilled employers which have been shown to be somewhat responsive to the minimum wage (Rohlin 2011 and Aaronson et al. 2018). Since our model treats the low and high-skilled as linked through production, we think of the high-skilled as the managers of those firms, with the very high-skilled not captured in our framework.

[15]If $G(0) < 0$, then local governments can set high minimum wages to induce the unemployed to migrate and raise total social welfare in the jurisdiction.

[16]When controlling for state-specific trends in the main results, the point estimate becomes small in magnitude and imprecise. The unweighted results are slightly larger in magnitude, but still imprecise. Although he interprets this as the minimum wage having no effect on high-skilled migration, the result is likely underpowered.

2.2.3 Extensions of the Decentralized Model

In our baseline model, the results and conditions for a binding optimal minimum wage are driven by the high-skilled migration externality: when a local government raises its minimum wage some high-skilled workers move out. The size of the externality depends on the choice of social welfare function and the production technology. We extend the model to see how changes in these assumptions alter the conclusions. We consider a social welfare function where governments only care about natives or maximize housing values, when there is a single firm that operates in both regions, and the introduction of additional inputs to production.

2.2.4 Social Welfare Function

In our framework, we assume that the government only cares about the total utility of its final residents in the second stage equilibrium. In the first part of this section, we consider a government that only cares about natives, regardless of whether they move.[17] When governments maximize utility of their natives, the derivative of the social welfare function shown in equation 2.2 no longer includes the final term: $\frac{\partial h_i}{\partial \bar{w}_i} G(c_h)$. However, they internalize migration costs. The government still trades off lower consumption of high-skilled workers with higher consumption of the low-skilled through the minimum wage, but no longer faces a first order cost from emigration. The implied minimum wages are higher compared to our baseline model.

Proposition 2. *If conditions 1-3 from Proposition 1 hold, the government cares only about their initial residents, and it values additional wages to low-skilled workers more than the loss of wages to high-skilled (redistribution) and moving costs, then it is total welfare improving for a jurisdiction to impose a (small) binding minimum wage.*

Proof. This is immediate from the first order condition, where the last term now concerns moving costs and not outmigration. Under this new assumption on the social welfare func-

[17] See Wilson (2015) for an overview of different assumptions on government maximization with resident mobility and Cremer and Pestieau (2004) for a review of the literature on government maximization with factor mobility.

tion, h_i is fixed.

$$0 < \frac{-\partial l_i}{\partial \bar{w}_i} G(0) + \frac{\partial l_i}{\partial \bar{w}_i} \int_{\Delta_{\bar{w}_i}\Theta_{iu}} \mu(\theta) G(c_l) d\theta$$
$$+ \int_{\Theta_{ie}(\bar{w}_i)} \mu(\theta) G'(c_l) d\theta \frac{\partial w_i^l}{\partial \bar{w}_i} + \int_{\xi(\Theta_{ih})} G'(c_h) \left(\frac{\partial w_i^h}{\partial \bar{w}_i} - \frac{\partial \xi}{\partial \bar{w}_i} \right) d\xi$$

□

Corollary 2. *Local governments that care about their initial residents set weakly higher minimum wage policies than those that care about their final residents.*

Proof. Comparing the first order conditions for a small binding minimum wage, this requires $\frac{\partial h_i}{\partial \bar{w}_i} G(c_h) \leq -\int_{\xi(\Theta_{ih})} G'(c_h) \frac{\partial \xi}{\partial \bar{w}_i} d\xi$. By construction, $\frac{\partial \xi}{\partial \bar{w}_i} = \xi$ if the worker moves and 0 otherwise, allowing us to rewrite the inequality as $\frac{\partial h_i}{\partial \bar{w}_i} G(c_h) \leq -\frac{\partial h_i}{\partial \bar{w}_i} G'(c_h) \mathbb{E}_{\Delta\xi(\Theta_{ih})}[\xi]$. First consider when $-i$ does not have a binding minimum wage. Then $\frac{\partial h_i}{\partial \bar{w}_i} < 0$, and the inequality holds since $G(c_h), G'(\cdot), \xi > 0$ by assumption. If $-i$ has a binding minimum wage and its initial residents return after i puts on a binding minimum wage, then both sides of the inequality are 0, and it holds as well. If i residents migrate instead, then the argument from the first case applies. □

The preceding analysis assumed that governments are utility maximizers, but we now consider a model where rent-seeking governments maximize property values.[18] In the most general application of this approach to our framework, all households and firms demand land. Define $H^d(\bar{w})$ to be the housing or land demand under minimum wage policy $\bar{w}$ and $\Delta_j^{j'} H^d(\bar{w})$ be the change for workers who transition from type j to j' due to the policy change, or for firm f. We assume that housing demand is nondecreasing in income, and that the elasticity of housing supply is finite. Then, for jurisdiction i, a binding minimum wage

[18]See Epple and Nechyba (2004) for a survey of these two different assumptions on the government's objective function.

is optimal if and only if

$$\frac{\partial H^d}{\partial \bar{w}} = \Delta_u^u H^d(\bar{w}) + \Delta_l^u H^d(\bar{w}) + \Delta_l^l H^d(\bar{w}) + \Delta_{h_s}^{h_s} H^d(\bar{w}) + \Delta_{h_s}^{h_s{}'} H^d(\bar{w}) + \Delta_f H^d(\bar{w}) > 0 \tag{2.3}$$

It is immediate that $\Delta_l^l H^d(\bar{w}) \geq 0$. Workers who remain employed after the wage increase will have more income and demand more housing. Similarly, workers that lose their jobs will demand weakly less housing, giving $\Delta_l^u H^d(\bar{w}) \leq 0$. The magnitude of this term importantly depends on how we treat the working cost. If θ is a real cost, then for a small binding minimum wage, the workers who lose their jobs under efficient rationing will have zero net income both before and after the policy change. If it is a utility cost, they will have less income and demand strictly less housing. In the welfare maximization framework, this distinction did not affect optimal policy. In the case where θ is a utility cost, the minimum wage is less attractive because we no longer have an assumption akin to efficient rationing. The always unemployed also have no change in income, implying $\Delta_u^u H^d(\bar{w}) = 0$.

Similar to our analysis in the utility maximization case, there is weakly less demand from high-skilled workers when the minimum wage increases. $\Delta_{h_s}^{h_s} H^d(\bar{w}) \leq 0$ since their wages weakly decrease and $\Delta_{h_s}^{h_s{}'} H^d(\bar{w}) \leq 0$ due to outmigration from falling wages. The changes are zero in the extreme case where low and high-skilled labor are perfect substitutes.

Finally, consider firms' demand for land. We can decompose the effect on firms as: $\Delta_f H^d(\bar{w}) = \frac{\partial H_f^d}{\partial L(l,h)} \frac{\partial L(l,h)}{\partial \bar{w}}$, where $L(l,h)$ is the total amount of land used by low and high-skilled workers for production. We showed that $\frac{\partial L(l,h)}{\partial \bar{w}} < 0$ as low-skilled workers become unemployed and the high-skilled migrate. If we assume that more workers requires more land, then $\frac{\partial H_f^d}{\partial L(l,h)} > 0$, and firm demand for land decreases. However, if the firm engages in more land-intensive activity when it has fewer workers, then $\frac{\partial H_f^d}{\partial L(l,h)} < 0$, and firm demand for land increases.[19] Summarizing the above analysis gives us the following proposition.

Proposition 3. *A rent-seeking local government will implement a small binding minimum*

[19] In Section 2.2.5 we derive the general effect of additional inputs on the optimal minimum wage.

wage if conditions 1-3 from Proposition 1 hold and if the housing supply elasticity is finite, housing demand is nondecreasing in after working-cost net income, and the net change in housing demand from the low-skilled employed and firms is positive and larger than the decrease in demand from the high-skilled.

Proof. Under the assumptions, the condition in equation 2.3 is satisfied as described above since the first two terms are 0, and the sum of the remaining are positive. Note that the final term can be either positive or negative. □

The optimal minimum wage trades off more housing demand from the still employed low-skilled with lower demand from the high-skilled due to wage decreases and outmigration. We discussed a comparable set of assumptions to the welfare maximizing case, (i.e. efficient rationing with no change in housing demand for the newly unemployed, and no land demand by firms), that leads to similar policies. We specify utility maximization in our main theoretical and quantitative analysis because the results are theoretically similar and it does not require us to also specify a specific housing market system. This allows us to focus our analysis on the high-skilled migration externality.

2.2.5 Capital and Other Inputs

With two inputs to production and zero-profits, high-skilled wages decrease with the minimum wage. We introduce elastically supplied capital, k, with required return, $\bar{r}$, so that this is not necessarily the case. With three production inputs, the zero-profit condition implies $\frac{dw^h}{d\bar{w}} = \frac{-l}{h} - \frac{k}{h}\frac{d\bar{r}}{d\bar{w}}$. If $\frac{d\bar{r}}{d\bar{w}}$ is relatively large in magnitude and negative, the marginal product of high-skilled workers increases with the minimum wage. The increased marginal product will induce migration into the jurisdiction as high-skilled workers move for higher wages. In this case, it is Pareto efficient[20] to implement a very small binding minimum wage since the marginal low-skilled worker is indifferent between working and not due to efficient rationing, and the wages of all other still employed low-skilled and all high-skilled workers increase.

[20] We are implicitly assuming that capital is paid its marginal product and always earns zero profit so we do not need to worry about rents to capital owners.

Raising the minimum wage further will not be Pareto efficient, and the government will trade off additional consumption from the employed low-skilled and high-skilled, and migration into the region from high-skilled workers with reduced consumption from the newly unemployed.

From the zero-profit condition, we can generalize $\frac{dw^h}{d\bar{w}}$ when there are additional inputs, indexed by ι with price p^ι. Then,

$$\frac{dw^h}{d\bar{w}} = \frac{-l}{h} - \sum_{\iota \neq h,l} \frac{\iota}{h} \frac{dp^\iota}{d\bar{w}} \tag{2.4}$$

Proposition 4. *When there are additional inputs to production besides high and low-skilled labor, a small binding minimum wage is Pareto Efficient if conditions 1-3 from Proposition 1 hold and* $\frac{dw^h}{d\bar{w}} \geq 0$

Proof. When $\frac{dw^h}{d\bar{w}} \geq 0$, there is no outmigration, and by efficient rationing all workers are weakly better off. □

2.3 Alternative Configurations for Minimum Wage Setting Authority

Using our baseline model, we consider the implications of alternative minimum wage setting authorities. As in the decentralized case, we develop conditions under which each type of government finds it optimal to implement a binding minimum wage.

2.3.1 Only Federal Government

In this section, we consider optimal minimum wage setting when only the federal government can set a minimum wage. The government's social welfare function is defined as

$$SW_f = \sum_i (1 - l_i)G(0) + \int_{\Theta_{ie}} \mu(\theta)G(c_l)d\theta + \int_{\xi(\Theta_{ih})} \zeta(\xi)G(c_h)d\xi \tag{2.5}$$

Note that in the federal government's optimization problem, it cares about the utility of all high skilled workers regardless of residence. Unlike in the regional government's problem, when high skilled workers move, their utility still matters for the federal government. The government also cares about the migration costs paid by individual workers.[21]

2.3.2 Federal Non-Uniform Policy

When the federal government can set a different policy in each region, the policies are given by:

$$(\bar{w}_1, \bar{w}_2) = \arg\max SW_f(\bar{w}_1, \bar{w}_2) \tag{2.6}$$

The FOC for $\bar{w}_1$ imply that it will be binding in jurisdiction 1 if:

$$\begin{aligned} 0 < & -\frac{\partial l_1}{\partial \bar{w}_1} G(0) + \frac{\partial l_1}{\partial \bar{w}_1} \int_{\Delta_{\bar{w}_1} \Theta_{1u}} \mu(\theta) G(c_l) d\theta \\ & -\frac{\partial l_2}{\partial \bar{w}_1} G(0) + \frac{\partial l_2}{\partial \bar{w}_1} \int_{\Delta_{\bar{w}_1} \Theta_{2u}} \mu(\theta) G(c_l) d\theta \\ & + \int_{\Theta_{1e}(\bar{w}_1)} \mu(\theta) G'(c_l^1) \frac{\partial w_l^1}{\partial \bar{w}_1} + \int_{\Theta_{2e}(\bar{w}_1)} \mu(\theta) G'(c_l^2) \frac{\partial w_l^2}{\partial \bar{w}_1} \\ & + G'(c_h) \left(\bar{h} \frac{\partial w^h}{\partial \bar{w}_1} - \int_{\xi(\Theta_{ih})} \frac{\partial \xi}{\partial \bar{w}_1} \zeta(\xi) d\xi \right) \end{aligned}$$

The expression above is simplified by noting that $\frac{\partial h_1}{\partial \bar{w}_1} = -\frac{\partial h_2}{\partial \bar{w}_1}; h_1 + h_2 = \bar{h};$ and $\frac{\partial w_h^1}{\partial \bar{w}_1} = \frac{\partial w_h^2}{\partial \bar{w}_1}$.[22] We also use the fact that moving costs are randomly assigned after households choose a residence in the competitive equilibrium and that under our assumptions, there will be net migration to jurisdiction 2. The FOC for $\bar{w}_2$ is defined analogously.

The first row of the FOC gives the welfare loss from newly unemployed low-skilled workers in jurisdiction 1. The second row gives the welfare gain from new low-skilled employment in jurisdiction 2. As high-skilled move to jurisdiction 2, the marginal product of low-skilled labor increases, leading workers to enter the labor market. The third row gives the welfare gain to the still-employed in both jurisdictions. The wage, conditional on employment, in-

[21] The final term of the government's social welfare function therefore involves integrating over the types of households in each jurisdiction.

[22] The last equality only holds if the lower bound on the moving costs are zero, which we maintain throughout, and for a small minimum wage.

creases in jurisdiction 1 due to the new binding minimum wage, while the wage in jurisdiction 2 increases because of induced high-skilled immigration. The final row gives the welfare loss to high-skilled workers, due to both the change in wages and moving costs paid by those who move to jurisdiction 2.

Proposition 5. *If (1) there is efficient rationing; (2) the demand elasticity is finite; (3) the supply elasticity of low-skilled workers is positive; (4) the government values additional wages to low-skilled workers more than the loss of wages to high-skilled; and (5) the first worker to move to the other jurisdiction has zero moving costs, $\xi = 0$, then it is total welfare improving for the federal government to impose a (small) binding minimum wage.*

Proposition 5 is the same as Proposition 1, except the central government only cares about migration insofar as it affects efficiency and moving costs paid. Condition (5) is similar to efficient rationing, but for moving, so that moving costs are zeros for a small binding minimum wage. It implies that the lower bound on the moving cost distribution is zero.[23] Even without condition (5), the federal government may implement a binding minimum wage if it values redistribution enough. An increase in the federal minimum wage leads to lower high-skilled wages, regardless of moving costs, and is therefore not Pareto improving. Comparing Propositions 1 and 5, we expect the federal government to prefer higher minimum wages than regional governments, all else equal.

The case in which the federal government can set a different minimum wage in each region is synonymous with the social optimum. The social optimum is defined in our context by the sum of total welfare for all agents in all regions. This is exactly the function maximized by the federal government. When its policy is allowed to be asymmetric, it has the flexibility to reach the socially optimal minimum wage in each region.

[23] This assumption on the moving cost distribution is not realistic and is chosen to provide sufficient, but not necessary, conditions for a federal minimum wage. However, in a long-run version of our model where workers do not have a fixed residence, the moving costs can be interpreted as preferences for one region compared to the other. In that case, it is plausible that some workers are indifferent, all else equal.

2.3.3 Federal Uniform Policy

In this case, we restrict the central government to set a single uniform policy. The minimum wage is given by:

$$\bar{w} = \arg\max SW_f(\bar{w}) \tag{2.7}$$

The FOC is quite similar to the non-uniform case, where the responses are given with respect to $\bar{w}$ instead of $(\bar{w}_1, \bar{w}_2)$. When $\gamma = 1$, the solution will be the same in the non-uniform and uniform cases. If $\gamma < 1$, then the competitive equilibrium low-skilled wages in jurisdiction 1 will be greater than those in 2. Therefore, when considering a small binding minimum wage, it will only be binding in 2. Under the assumptions of Proposition 5, the Social Planner will always want a binding minimum wage in each jurisdiction, however it may be the case that even when the assumptions hold, the federal government does not set a binding minimum wage if constrained to only set a uniform policy. When the Social Planner sets a binding policy in each jurisdiction, it can mitigate the welfare loss due to migration, both moving costs and the suboptimal allocation of individuals across jurisdictions. When the federal government can only set one policy and it is only binding in a jurisdiction, then there is an externality imposed on the other. This is seen by noting that all of the components of the government's first order conditions depend on the enforced minimum wage in both jurisdictions. For $\gamma < 1$, a small binding uniform policy is equivalent to the Social Planner only setting a minimum wage in the low productivity region, which the previous section showed is suboptimal.

Tiered System

The US has a tiered system in which the federal and local government both set minimum wages, and the larger of the two is enforced. The federal government is restricted to set a uniform policy. Based on our previous analysis, under the assumptions of Proposition 1, each local government sets a binding minimum wage, and under the assumptions of Proposition

5, the federal government also sets a binding policy for at least the less productive region. As a result, there will be binding policy enacted in each. In the symmetric case, the federal minimum wage will be weakly higher than both jurisdictions' preferred policies because of the differential effects of outmigration and moving costs. The effect of moving costs must be less than outmigration by construction.

2.4 A Two Region Example

2.4.1 Parameterization

In the interest of exploring the relative welfare consequences of different configurations and the relative advantages of decentralized and centralized policy setting, we consider a parameterized version of the model. The calibrated model illustrates the mechanisms highlighted in the theory section, particularly regional heterogeneity and competition. In the previous sections, we are unable to give theoretical predictions for each policymaker beyond the extensive margin conditional on the other policies. The calibrated simulations in the next section on optimal policy provide results for the intensive margin decisions when governments compete while factoring in the general equilibrium effects of their actions.

We specify the following functional forms. We assume f is a constant elasticity of substitution (CES) production function with elasticity of substitution between high and low-skilled labor $\sigma = \frac{1}{1-\rho}$ and weight α on the low-skilled.

$$f(l,h) = (\alpha l^{\rho} + (1-\alpha)h^{\rho})^{\frac{1}{\rho}} \tag{2.8}$$

With this production function, $\rho \leq 1$. If $\rho = 1$ the two types of labor are perfect substitutes and if $\rho = -\infty$ then the two types are perfect complements. Worker heterogeneity is determined by the distributions on θ and ξ, which we assume to be $\theta \sim \mathcal{U}(0,\bar{\theta})$ and $\xi \sim \mathcal{U}(0,\bar{\xi})$. The sufficient conditions from propositions 1 and 5 require that the lower bounds are zero. Alternative distributional assumption on the moving costs do not qualitatively affect our results.[24]

We also assume that the social welfare aggregator is defined as:

$$G(c_j) = \int_{j \in r} \log(c_j + \nu) dj \tag{2.9}$$

where ν is a utility function shifter. Since some low-skilled workers are always unemployed with $c_u = 0$, ν must be positive. $\log(\nu)$ is therefore the value of unemployment to the government. Smaller values of ν do not affect labor supply, but increase the social cost of unemployment. We choose individual utility to be net, after migration and disutility from working, consumption so that there are no income effects.

2.4.2 Calibration

We calibrate our model to match aggregate moments of the US economy. We have seven parameters: α, γ, and ρ for the production function; $\bar{h}, \bar{\theta}$, and $\bar{\xi}$ to describe the worker population; and ν for the social welfare function. We vary γ to see how different levels of regional heterogeneity impacts our results; for the calibration, we set $\gamma = 1$, corresponding to the symmetric case. Additionally, we let $\bar{h} = 1$ to reflect the fact that about one-third of the population has a Bachelor's degree (CPS), and $\rho = 0.286$, based on the estimate of the elasticity of substitution between high and low-skilled workers in Katz and Murphy (1992).[25]

Based on the timing assumptions of our model, we calibrate the remaining parameters in two stages. In the first stage of the calibration, based on the competitive equilibrium, we set α and $\bar{\theta}$. We calibrate α to match the relative wages across skill types. Using the Current Population Survey Outgoing Rotation Group microdata, the Economic Policy

[24] For example, modeling the moving costs as a type I extreme value idiosyncratic preference for region 1, similar to our quantitative model in Section 2.6, does not affect the mechanisms highlighted here. In both cases, low ξ high-skilled workers in i will move after an increase in the minimum wage in i. However, the number of movers, and therefore optimal policy, depends on the distributional assumptions. We specify that the moving costs are normally distributed so that all draws of ξ are non-negative for all values of γ. Under the type I extreme value assumption with asymmetric regions, some individuals have negative moving costs in the baseline. For the purposes of this exercise, we do not allow for that source of misallocation.

[25] For the purposes of the calibration, we set $\bar{h}$ using college graduates to match the definition of high-skilled in Katz and Murphy.

Initiative reports hourly wages in 2017 by highest degree received. Based on their report, we set $\frac{w^h}{w^l} = 1.82$. We calibrate $\bar{\theta}$ to match the 2017 labor force participation rate of high school graduates. Based on estimates from the Bureau of Labor Statistics in 2017, the value for high school graduates is 57.7%, while the value for those with only a Bachelor's degree is 73.3%. Since our model assumes that high skilled workers are always in the labor force, we scale up the high school rate to $\frac{57.7}{73.3} = 78.7$. ν is normalized to 1 so that the utility of the unemployed is 0. It also ensures that $G(c_h) > 0$

In the second stage, after all workers have established residencies from the competitive equilibrium, we calibrate the upper bound on the moving cost distribution, $\bar{\xi}$ to match the elasticity of migration with respect to wages of 0.5 from Kennan and Walker (2011). For a 10 percent decrease in γ, $\bar{\xi}$ is found so that there is an increase in the stock of high-skilled workers in jurisdiction 1 of 5 percent.[26]

The parameter estimates are presented in Table 2.1. The estimate column gives the values used in the welfare calculations and the residuals note the distance between the model implied value and the target, when relevant.

From the baseline calibrated model, Figure 2.1 plots total social welfare in the economy as a function of the minimum wage. In the left panel, the minimum wage is effective in both regions and so this figure represents the problem of the central government restricted to a uniform policy. A small binding minimum wage increases welfare since the government values redistribution, the marginal separated low-skilled worker has 0 surplus from working, and no high-skilled workers move since the regions are identical and a uniform policy is enforced. As the minimum wage increase further to 102.5 percent of the competitive wage, welfare also increases. After this point, the higher price floor causes relatively too much additional unemployment and lost high-skilled wages. When the minimum wage is 5.6 percent above

[26]We interpret the Kennan and Walker (2011) partial equilibrium results as a productivity shock in our general equilibrium model. The shock in one region affects high-skilled wages directly and through its effect on low-skilled labor supply, while the general equilibrium migration effect moves in the opposite direction. If instead we shock just high-skilled wages in one state holding everything else fixed, then $\bar{\xi}$ would barely change to 1.3140 instead of 1.3134.

the competitive level, the policy is worse than no policy at all.

The right panel of the figure plots total welfare as a function of the minimum wage in region 1 with no minimum wage in the other. If region 2's government does not set a binding policy, this figure depicts the decentralized problem for region 1. For a small binding policy, total social welfare decreases in region 1 because the loss of high-skilled workers is more than the value from additional wages to the low-skilled. Welfare in region 2 increases from the inflow of new high-skilled workers, which increases the total population and low-skilled wages but at the expense of high-skilled wages. The increase in region 2 is not enough to offset the loss in region 1 and so a small binding minimum wage in only one region decreases total welfare in the economy. The migration externality is too large relatively to the benefits from redistribution.

2.5 Optimal Minimum Wage Setting

To supplement the theoretical work on the extensive margin, we use the calibrated model to trace out the relative benefits of centralization verse decentralization as a function of mobility and regional heterogeneity for optimal minimum wage setting. Our results illustrate the trade-off captured by the decentralization theorem, where the federal government can only set a single uniform policy and local governments can better adjust to their own needs. We find that centralization of minimum wage setting authority is always weakly preferred to decentralization in our baseline calibrated model, even when regions are heterogeneous, because of interregional spillovers and the corresponding competition for high-skilled workers. Local governments never want to set binding minimum wages, while the central government loses its incentive to set a uniform minimum wage as heterogeneity across regions increases.

Panel A of Table 2.2 presents results for the symmetric case, i.e., when relative productivity $\gamma = 1$. The five rows correspond to the five different types of government structure that we allow in our context. The top row is the competitive equilibrium with no government, and the bottom row is the case of a benevolent central government setting non-uniform policy, which is synonymous with a social planner solution. The middle three rows correspond to cases with local governments only, a central government restricted to uniform policy only, and a US-style hybrid system with both of these types, respectively. Columns list the values for unemployment, wages for high-skilled and low-skilled workers, and the fraction of (mass 1) high skill workers that sort to region 1. If the wages of a particular case match the wages in the competitive case, then there is no binding minimum wage. The far right column lists the welfare loss relative to the federal non-uniform or social planner.

In the symmetric case, local governments do not want to set a binding minimum wage because the resulting outmigration is too large to be offset by the increase in wages for the low-skilled. Condition 4 of Proposition 1 does not hold. In contrast, the central government is not affected by this race-to-the-bottom behavior to keep high-skilled workers because migration does not change the set of agents in the central governments welfare function. In the symmetric case, the central government exactly matches the social planners policy, since the social planner sets identical minimum wages in each region.

In our calibrated equilibrium, we find that the social planner and central government set the optimal minimum wage to 102.5% of the competitive wage. The binding minimum wage leads to a small total welfare increase over the competitive and decentralized cases of 0.23%. With higher low-skilled wages, the always employed low-skilled have a welfare gain at least 2.14% that depends on their working costs. The gain comes at the expense of the newly unemployed, whose welfare decreases by 100% as consumption drops to 0, and the high-skilled, whose welfare falls by 1.65%. As more low-skilled workers become unemployed, the marginal product of the high-skilled workers falls, leading to lower consumption. The always unemployed are unaffected by changed in the policy.

With heterogeneous regions, the central governments ability to match the social planner's solution disappears. Panel B of Table 2.2 presents the outcomes when γ is reduced to .995. With this very small productivity difference, the central government desires a binding minimum wage that is between the social planners two policies: slightly too high for the less productive region, but too low for the more productive area. The centralized and tiered environments both result in a positive welfare loss relative to the social planner, though they are still preferred to the decentralized case where local governments do not implement a binding policy.

Finally, we present results for $\gamma = 0.99$ in Panel C of Table 2.2. In this case, the local and central uniform governments do set a binding minimum wage, while the social planner's solution sets a different binding policy in each region. As in the previous cases, the migration incentive is too strong for the local governments to set a binding policy, and now that is also the case for the central uniform as well. A small binding central uniform policy only affects the less productive jurisdiction, which causes high-skilled workers to migrate. The migration externality decreases welfare as total output in the economy shrinks, even more low-skilled workers in region 2 become unemployed, and some high-skilled pay migration costs. In the $\gamma = 0.995$ case, the jurisdictions were similar enough that the same policy implemented in both did not cause too large of an externality. With large amounts of heterogeneity, like in the US, central uniform policies are inefficient. The optimality of a central uniform minimum wage is decreasing in regional heterogeneity. With two policies to set, the social planner is able to set the binding minimum wage in each jurisdiction so that the populations remain the same as in the competitive equilibrium.

2.5.1 Variable Mobility

The relative differences in government competition of mobile high-skilled agents drives our main results in the previous section. Local governments are averse to minimum wage induced outmigration, as it creates a first order welfare loss. The central government faces a much smaller cost from migration, especially since the first workers who move have no migration costs. However, both types of governments bear the burden of decreased total output as households sort across jurisdictions. We now explicitly vary the mobility of the high-skilled to better understand this mechanism. With high migration costs, the model captures the short-run effects as workers are unable to move in response to the policy. As migration costs decrease, the model better captures the long-run equilibrium effects.

When migration costs increase, minimum wages become more desirable for regional governments; when high-skilled workers are immobile, regional governments enact exactly the social planner policy. Without migration, there are no interregional spillovers and the local governments problems exactly match that of the social planner because the two economies are additively separable in the social welfare function. From the baseline symmetric model, if the minimum of the moving cost distribution is increased slightly to 0.0025 from 0, or 0.3% of the competitive high-skilled wage, then the decentralized outcome would match the social planner as well. The lowest moving cost is large enough that the high-skilled are effectively immobile in equilibrium. As long as the minimum moving cost is more than 0.03% of the high-skilled competitive wage, regional governments will enact binding policy in the baseline model. Small changes in the moving cost have large implications for government competition.

To understand the interaction between productivity heterogeneity and mobility, we set γ to .95 and the minimum of the migration cost distribution to 0.0025. Table 2.3 presents the results. As before, the social planner implements a different binding minimum wage in each jurisdiction. In the decentralized case, migration costs lessen horizontal competition enough so that both local governments set binding policies, but not enough to replicate the social planner's solution. The central uniform policy alone leads to greater welfare loss than the decentralized policy setting because there is considerable heterogeneity, but the government only has one instrument. However, the hybrid system captures the social optimum. When the central government sets its policy equal to that of the social planner in jurisdiction 2, it now becomes optimal for jurisdiction 1 to raise its policy. The local government faces a smaller threat from outmigration than the decentralized case where the regional government in jurisdiction 2 sets its policy too low. In this more realistic case, where jurisdictions are heterogeneous and all migration is at least a little costly, the hybrid minimum wage setting authority is welfare improving over either by itself. Centralized and decentralized policy setting are strategic complements. This justifies the policies we see in Seattle, for example, while not supporting the policies by the Missouri state government of prohibiting its cities from setting higher minimum wages or in Europe where many countries set a single national policy.

At the other extreme, we consider perfectly mobile high-skilled workers. In this case, the outmigration response to a minimum wage will be even stronger than in the baseline case, such that local governments will have even less of an incentive to implement binding minimum wages. For the central government, the effect is ambiguous. The lack of moving costs means that migration is less costly, since what matters to the central government is the costs paid by individuals rather than the loss of population. However, the central government also cares about the total output in the economy, which is decreasing with migration. When $\gamma = 1$, no households move as a result of federal policy, and the calibrated migration and perfect mobility equilibria are identical. When $\gamma = 0.95$, the federal non-uniform optimal policies are 0.379 in jurisdiction 1 and 0.340 in jurisdiction 2, compared to 0.383 and 0.343 in the baseline calibrated case. Higher levels of mobility push down optimal federal policy.

2.5.2 Alternative Social Welfare Functions

The previous quantitative results rely on a single specification of the social welfare function, namely, $U(c) = \log(c+\nu)$, where we set $\nu = 1$. ν is a shifter on the progressivity of the welfare function, with lower values corresponding to a more progressive desire for redistribution. The right panel of Figure 2.2 presents the equilibrium minimum wages desired by local governments and by the social planner as a function of the ν for the symmetric case holding all other structural parameters fixed. Since our theory requires that $G(c_h) > 0$, ν cannot be too small in our baseline model. To fully investigate the importance of progressivity, the left panel of the figure uses the social welfare function $\log(2(c+\nu))$ so that $G(c_h) > 0$ for all $\nu > 0$ since in the symmetric baseline equilibrium, $c_h > .6$ for any feasible minimum wage policy.

The optimal minimum wages are an inverted-U shape in progressivity. As ν approaches zero, the welfare of the unemployed quickly heads toward negative infinity, such that adding to unemployment becomes increasingly costly from a social welfare perspective. Minimum wages distribute from the high-skilled and the newly unemployed to the low-skilled employed. If ν were 0, then the social welfare would be negative infinity for all policies, including no binding minimum wage, since there will always be some unemployed workers in our model. In that case, the optimal minimum wage would not be defined for any government. The pattern is identical when $\gamma < 1$.

We now consider a utilitarian social welfare function. Since our agents maximize consumption, this social welfare function is equivalent to maximizing the sum of wages minus any working costs or moving costs paid. In the symmetric $\gamma = 1$ case, no binding minimum wages are optimal for any government. With a relatively large amount of regional heterogeneity, for example when $\gamma = 0.95$, the social planner desires a small binding minimum wage of about 0.3% over the competitive wage in each jurisdiction. Under our baseline social welfare function, when $\gamma = 0.95$, the social planner would set optimal minimum wages of about 2.5% above the competitive.

In addition to the productivity heterogeneity explored previously, regions may also differ in their progressivity. St. Louis and Kansas City, Missouri as blue cities in a red state may have set higher minimum wages because they value redistribution more. To test this implication, Figure 2.3 varies ν for only jurisdiction 1, holding the value for jurisdiction 2 fixed at 1, and plots the equilibrium decentralized policies and the competitive outcome when jurisdictions are otherwise identical. When ν is slightly less than 1, jurisdiction 1 values redistribution sufficiently to set a binding minimum wage. Unlike in the baseline, condition 4 of Proposition 1 is now satisfied. Although the welfare cost of setting a minimum wage for the government in jurisdiction 2 is now lower because the binding minimum wage in the other region reduces the value of migration for the high-skilled, it remains at the corner of no binding policy. It is not until ν in jurisdiction 1 falls below 0.7 that jurisdiction 2 sets a binding minimum wage. As ν in jurisdiction 1 falls further, jurisdiction 2 is induced to set a binding minimum wage. In this region of ν, governments engage in race to the top. When

ν decreases even further, the benefit of setting a high minimum wage also decreases and the policies begin to level off. Even when regions are equally productive, differences in the value of redistribution can lead competing governments to set binding decentralized policies.

2.5.3 Information

In addition to mobility and regional competition, the relative information quality of different levels of government affects the optimal minimum wage setting authority. Decentralized policy setting may be preferred because local governments have better knowledge of the labor supply, labor demand, and migration elasticities or local productivity. These parameters jointly determine how employment changes after a minimum wage increase. In a 2019 report, the CBO documented a large amount of variation in estimates of the employment elasticities from the recent minimum wage literature from 0.4 in Cengiz et al. (2019) to -1.7 in Clemens and Wither (2016). Which estimate should a government use when determining its policy? In the context of our model, if a government thought the employment elasticity were positive, then it would believe that a binding minimum wage is Pareto improving and only faces a welfare trade-off when the estimate is negative. Therefore different estimates lead to very different "optimal" policies. Local governments may additionally have better information about the market structure, which we do not explore in this paper, or political beliefs leading them to favor one estimate over another; for example, if it believes there is significant monopsony power in the region, it may place more weight on positive estimates. If any government has noisy or incorrect information about the regions or these parameters, then it will not be able to set policy optimally.

While the previous analysis assumed that both levels had perfect information about workers and regions, we now explore the consequences of misspecified policy.[27] Panel (a) of Figure 2.4 presents the welfare for different minimum wage policies in each region for the baseline model normalized by the welfare in the competitive case with no binding policy.

[27] We do not make explicit assumptions about why the minimum wage differs from the optimal policies found in the previous subsection but instead simulate how deviations affect welfare. These errors may be due to a lack of information.

In each row, the minimum wage varies from 0 to 10 percent above the competitive wage in region 1. The columns vary the price floor in region 2. The grid therefore constructs a discretized version of the social planner's problem and it confirms that the optimal minimum wage is about 2.5 percent above the competitive. Similarly, looking down the main diagonal gives the optimal central uniform policy. Based on these simulations, we see that very large uniform minimum wages can lead to lower total welfare than having no policy, which is also the decentralized case. Importantly for policy-making, these results suggest that when regions are symmetric, a minimum wage that is too high is better than one that is too low. Specifically, we find that a minimum wage 0.5 percentage points too high yields higher welfare than one that is 0.5 percentage points too low. The same is true when we compare policies that are 1.5 and 2.5 percentage points off in either direction.[28]

Asymmetric policies for which the difference is more than 2 percentage points also lead to lower welfare than having no policies at all. When the enforced minimum wage is too different in otherwise identical regions, the migration externality overpowers the benefits from redistribution. Since symmetric policies do not induce migration, relatively more of the burden falls on high-skilled workers that the progressive government values less. However, large differences in the enforced low-skilled wage across equally productive regions leads to a large amount of migration. This exacerbates the disemployment effects of the minimum wage in the region with a higher policy. In our simulations, we find that minimum wages of 101.5 and 103.5 percent of the competitive performs worse than a uniform policy of 106 percent.

To understand how errors in optimal policy setting from a lack of information vary with productivity and mobility, Panel (b) displays the relative welfare based on the simulations shown in Table 2.3 where $\gamma = .95$ and $\min(\xi) = 0.0025$. With perfect information, the optimal uniform central policy is less than 101 percent of the competitive wage in region 2 and not binding in region 1 while the optimal non-uniform policy is about 102.5 percent

[28]We cannot compare more than 2.5 percentage points off of the optimal because then the low minimum wage would not be binding.

above each region's respective competitive wage.[29] Since moving costs are positive on the margin, both regions set binding policies of about 1 percent above the competitive in the decentralized case as well. The patterns are very similar to the baseline example except the welfare matrix is no longer symmetric. Based on the asymmetry, we want to determine if the government should be relatively more cautious in the high or low productivity region. Our results indicate that setting too high a minimum wage in the low-productivity region while setting too low a minimum wage by the same percentage points in the high-productivity region is better than the reverse.

2.5.4 Taxes and Transfers

In the previous sections, we study the optimal minimum wage without taxes, which is an “optimal suboptimal policy. When the federal government has access to linear income taxes (a single income tax rate and a demogrant), it will always prefer linear taxes to a combination of a minimum wage and linear taxes in our calibrated model. The proof of this numerical result in our framework is presented in Appendix B.1. Lee and Saez (2012) prove in a model without migration, when workers are not mobile across skill levels and non-linear taxes and transfers are available, then it is not optimal to also have a binding minimum wage. This is analogous to the no mobility case of our model discussed above. In a search-and-matching framework, Lavecchia (2018) finds that both a minimum wage and optimal taxes are preferred only if the government has very strong redistributive preferences.[30] As shown in the previous sections, the migration externality from any policy that restricts low-skilled labor supply and induces migration is worse for the local governments than the federal, and income taxes are no exception. This is intuitive in light of the literature which tends to find

[29]The diagonal of this figure does not give the optimal central uniform policy because regions have different competitive wages. Since productivity in region 2 is much lower than 1, the optimal uniform minimum wage is found by looking across the top row.

[30]Redistributive policies may also have an insurance component that high-skilled workers care about (Hoynes and Luttmer 2012). If a high-skilled worker was separated from her job due to a shock, and then was able to find employment at the minimum wage rather than the low-skilled competitive one, the insurance value would be positive. However, in our static model with fixed worker skill levels, the high-skilled are made weakly worse off from a minimum wage increase.

that redistribution is better handled by the federal government (e.g. see Gordon and Cullen 2012).

If the government taxes income, then low-skilled workers will restrict their labor supply, impacting the marginal product of all workers and lowering total output in the economy. High-skilled labor supply is not perfectly inelastic to each jurisdiction, but it is perfectly inelastic to the economy and the federal government. The government trades off redistribution with efficiency. A binding minimum wage will increase low-skill labor supply, but decrease labor demand, while an income tax will decrease labor supply and have no effect on demand, potentially allowing the policies to be complements. While the optimality of jointly specifying a minimum wage and linear income tax is uncertain, we show that under the assumptions of our calibrated model, the federal government would never implement a binding minimum wage if linear taxes are available.

Although it is not optimal to have both a binding minimum wage and a federal income tax, we incorporate this into our model and recompute the equilibrium to compare the policies. Our simulations confirm that the government does not want to set a binding minimum wage when it sets the income tax optimally. In our baseline symmetric calibrated model, we find that federal government would choose to set a tax rate of 24% and no binding minimum wage in both jurisdictions, as expected. The policy dramatically increases unemployment to 35.5% of the low-skilled workers from 26.2% under the optimal minimum wage. The restricted labor supply leads low-skilled wages to rise to 0.390, and high-skilled wages to fall to 0.616. The government trades off its redistribution incentive with decreased low-skilled labor supply that leads to increased unemployment and falling high-skilled wages. As the income tax increases, the marginal (zero consumption) low skilled worker exits the labor force. This raises the marginal product of all the still employed low-skilled workers, thus acting very much like a minimum wage, except now the government is also able to redistribute some of the additional income to the unemployed.

The optimal linear income tax increases total welfare over the optimal minimum wage by

2%. Compared with the optimal minimum wage model, where the utility of the unemployed is 0 by construction, the utility of unemployed is 0.126. The low-skilled worker with the lowest working cost sees her utility increase by 16.2% under the optimal linear income tax compared with a minimum wage, which is a lower bound on the welfare increase for low-skilled workers. High-skilled workers see their welfare fall by 5.1%, as the government redistributes total wealth. If the federal government were able to set a different tax rate on each skill type, it would tax low-skilled at 73% and the high-skilled at 100%, leading to a welfare gain of 8.5% over the optimal minimum wage. This result is driven by the fact that the high-skilled workers supply labor inelastically to the economy. Since the government can use that revenue to lessen the harm of unemployment, it then sets a much higher tax rate on the low-skilled as well.

Throughout this section, we assumed that the governments tax income. However, if the tax is on consumption, equal to income minus cost of working and migration, then there is no labor supply response, and a federal government with a concave social welfare function would optimally set a leveling tax so that all workers have the same utility. With a leveling tax, a minimum wage is suboptimal as the induced unemployment lowers total output in the economy that can be redistributed. A utilitarian government would have no desire to redistribute through a consumption tax since the total output consumed would not change.

2.6 A Quantitative Model of Minimum Wage Setting

Our two region model demonstrated that decentralized and centralized minimum wage setting are strategic complements and that the extent to which they are depends on regional heterogeneity and mobility. While this approach highlights important mechanisms, we now extend the model to the case of US states to quantify the welfare gain from the tiered policy compared to other minimum wage setting authorities.

2.6.1 Model

We consider a spatial equilibrium model of local labor markets as in Rosen (1979) and Roback (1982) with two imperfectly substitutable types of labor, and state and federal governments. High-skilled workers in state j have indirect utility given by

$$V_j^h = w_j^h + \phi_j + \kappa\epsilon_j \tag{2.10}$$

where w_j^h is the high-skilled wage rate in j, ϕ_j is a state j parameter that captures amenities and costs of living, and ϵ_j is an i.i.d. type I extreme value idiosyncratic preference for j, with standard deviation κ. High-skilled workers pick the location that maximizes their utility, which implies that the probability a worker lives in j is

$$p_j = \frac{\exp\left(\frac{w_j^h+\phi_j}{\kappa}\right)}{\sum_m \exp\left(\frac{w_m^h+\phi_m}{\kappa}\right)} \tag{2.11}$$

Immobile low-skilled workers can be employed or unemployed. If a worker i is employed in state j, she pays working cost $\theta_{ij} \sim \mathcal{U}(0,\bar{\theta}_j)$ and receives wages w_j^l. Therefore, her indirect utility is

$$V_j^l = \begin{cases} \phi_j & \text{if unemployed} \\ w_j^l - \theta_{ij} + \phi_j & \text{if employed} \end{cases} \tag{2.12}$$

Low and high-skilled workers in the same state face the same cost of living and receive the same amenities.

Each state has one firm that hires low and high-skilled workers with production function

$$f_j(l,h) = \gamma_j(\alpha_j l^\rho + (1-\alpha_j)h^\rho)^{1/\rho} \tag{2.13}$$

This specification differs slightly from Equation 2.8 in that it allows both γ and α to vary across markets. The government's social welfare function is given in Equation 2.9.

2.6.2 Data and Calibration

We calibrate the model to the continental US states in 2015. In 2013 and 2014, Congress considered the Minimum Wage Fairness Act, which would have raised the federal price floor in 2015 for the first time since 2009. After this failed, President Obama continued to push for an increase through 2015, although Congress did not act. Since the federal government considered changing the minimum wage, but chose not to, we take the 2015 policy to accurately reflect the government's preferences for redistribution. This allows us to back out ν for the federal government from the observed equilibrium.[31] While there has been a continued discussion of increasing the minimum wage, the federal government has not considered changing it since. We therefore use the 2015 American Community Survey (ACS) to estimate the model parameters.

From the ACS, we observe individual level wages by education and labor force attachment in each state. This gives the number of the low (high school equivalent) and high-skilled (college graduates) employed workers, as well as their average wages, the number of unemployed, and the number not in the labor force. Since many high school equivalent workers earn more than the minimum wage, the model implied high-skilled wage equals the state's minimum wage times the high-skilled/low-skilled wage ratio in the data. $\bar{\theta}_j$ in each state is picked to match the state's high school equivalent labor force participation rate in the data.

Taking log of Equation 2.11, we first estimate κ given the observed population shares and high-skilled wages in the data from the following regression equation:

$$\log(p_j) = \beta_0 + \beta_1 w_j^h + v_j \tag{2.14}$$

[31] More specifically, ν is calibrated so that \$7.25 is the welfare maximizing central minimum wage policy, holding the state policies fixed at their 2015 levels. This implies a Nash equilibrium where the federal government does not find it beneficial to deviate given the strategies of the states. However, the governments may be more sophisticated and know that changes in federal policy will also impact the state's optimal policies. We address calibrating ν for all states in Section 2.6.5.

where $\beta_1 = 1/\kappa$. However, endogeneity from the correlation of amenities and wages as well as measurement error will lead to a biased estimate of κ. We therefore instrument for the high-skilled wage with the number of low-skilled residents. The model provides intuition for it as a valid instrument. First, the marginal product of the high-skilled depends on the number of low-skilled employed in a state, and so it should be correlated with the total stock of low-skilled. We confirm this in the middle column of Table 2.4. Additionally, low-skilled workers are immobile, and so the total number is exogenous in the model. We find that κ is 1.048. As expected, the estimate is slightly larger than the value of 0.717 from Surez Serrato and Zidar (2016), who conduct a similar analysis on county groups, a smaller unit of geography, using a Bartik instrument approach. They similarly find an OLS/IV ratio for κ of just above 3.

After estimating κ, we can recover ϕ_j up to a normalization that $\min \phi_j = 0$. Since workers' indirect utility is a function of ϕ_j, the normalization together with our estimate of ν determine the concavity of the social welfare function. Under the functional form assumptions, this does not affect high-skilled location choices and is chosen so that all workers have non-negative utility. Since only $\phi_j + \nu$ is identified, an alternative normalization would only affect the magnitude of ν and not the results.

The calibration of the state specific production function follows Section 2.4.2. First, we set $\rho = 0.286$ following Katz and Murphy (1992). We calculate α_j to match the wage gap between high and low-skilled workers and γ_j so the marginal product of the low-skilled employed matches the state's minimum wage. Given the model's parametrization, the values perfectly match the targeted moments. [32]

After recovering the other parameters, we finally back out the federal government's redistribution preferences ν so that the observed policy of \$7.25 maximizes total social welfare in the economy. We aim to minimize $|7.25 - w^*(\nu)|$, where $w^*(\nu)$ is the optimal minimum

[32] For this calculation, we calculate the state's wage gap as the ratio of college graduates and high school equivalent or less in the ACS for employed workers between ages 16 and 70. Since the minimum wage is a floor, the ratio of high-skilled wages to the minimum wage would overstate their relative marginal product in the model.

wage as a function of the concavity of the social welfare function. Table 2.5 summarizes the parameter estimates. The model is unable to perfectly match the current federal minimum wage for any progressivity shifter, although the residual is only about 1 cent. In the context of the model where states are heterogeneous, the federal government and 21 states should not coordinate on the same policy. However, \$7.25 may be salient at a national level as a reference point, or the federal government may understate regional heterogeneity because of a lack of information.

2.6.3 Welfare and US Minimum Wage Setting Authority

Does joint state and the federal tiered minimum wage setting in the US increase aggregate social welfare over centralized policy alone? While our previous theoretical results suggest there should be positive gains, in the application, the sign is ambiguous since state are heterogeneous in their preferences for redistribution as well as their productivities. This implies that the most productive states do not always set the highest minimum wages. The correlation between the 2015 state policies and γ_j is only 0.53.[33] If a relatively low productivity state sets too high a minimum wage relative to what is binding in others, too many high-skill workers move. Since the state is of relatively low productivity, this disproportionately harms its low-skilled workers by increasing unemployment. Given this tension created from differences in preferences for redistribution, it is possible for optimal centralized policy to lead to higher total welfare from the perspective of the federal government than the tiered system.

To measure the potential gain or loss, Panel (a) of Figure 2.5 plots total welfare in the economy under both minimum wage setting authorities. The solid line shows how welfare in the tiered system changes as a function of the federal minimum wage. Based on the calibration, when the states' minimum wages are held at their 2015 levels, the federal gov-

[33] This calculation may overstate heterogeneity in preferences because states whose optimal policies are below \$7.25 have no incentive to set their own. Using the 2015 effective minimum wages, the correlation is 0.85, but for similar reasons this likely overstates heterogeneity. Heterogeneity in other production parameters matters as well. The correlation between α_j and the state policies is -0.36.

ernment optimally sets a minimum wage of nearly \$7.25, which confirms our calibration of ν is successful. The dashed line in the figure shows welfare as a function of centralized policy. Without the state policies, the federal government optimally chooses a lower minimum wage of \$6.90. Panel (a) of Figure 2.6 shows the relationship between tiered minimum wages and low-skilled wages under centralization.[34] This result is in line with the previous findings that state and federal polices are strategic complements. Higher state minimum wages reduce the high-skilled migration externality and allow the federal government to engage in more redistribution by setting a higher wage floor itself.

Optimal centralized policy has a welfare loss of about 0.006 percent compared to the tiered system.[35] Although the aggregate welfare change is small, state and federal minimum wages together are better able to redistribute to the low-skilled employed, as shown in Panel (b) of Figure 2.5. The figures also show that suboptimal policy in the form of low wages, for example due to information frictions discussed earlier, has a higher social welfare cost under tiered policy setting. This is because many states set their own policies at or above the federal level. High state policies become more costly when the federal price floor falls. The opposite is true when the federal governments sets too high of a minimum wage.[36] However, as the federal policy increases, the tiered and central uniform systems become more similar since only the higher of the state and national policy is implemented. If the federal policy were above \$9.47, the two would be equivalent.

While the current US minimum wage setting authority yields higher welfare than centralized policy alone, it may be far from the social planner's solution. The 48 states are heterogeneous in their productivity and low-skilled working costs, but 21 have effective minimum wage policies of \$7.25. This is in part because states face a different trade-off than the federal government as emphasized in equations 2.2 and 2.5, and because some states may be

[34]Based on our calibration, the tiered minimum wages are binding in all states, while our simulations predict that the optimal federal uniform minimum wage is only binding in 17 states.

[35]The competitive equilibrium has a welfare loss of 0.021 percent.

[36]When interpreting the value of the tiered system across different minimum wage levels, it is important to note that the analysis in this section holds state policies fixed. If the federal minimum wage were to fall (rise), states should lower (raise) theirs as well.

more or less progressive. Panel (b) of Figure 2.6 shows the relationship between the observed tiered policies, and those under federal non-uniform. States below the 45 degree line have too high minimum wages under tiered policy setting. When not limited to a single policy, the government can better differentiate among states. Those subject to the \$7.25 minimum wage in the tiered system have federal non-uniform policies that range from \$7.14 to \$7.53. The federal government also sets slightly higher minimum wages for states with initially high policies in this case since it is able to better internalize the spillovers from high-skilled migration. Overall, current US policy closely matches the social optimum predicted in this model with only a 0.05 percent welfare loss.

2.6.4 State Heterogeneity and Regional Redistribution

The minimum wage setting authority potentially impacts redistribution across states in addition to skill groups. Under centralization, the federal government has one policy lever that only directly affects the lowest productivity states. Higher federal uniform policies will therefore redistribute to more productive states. To some extent, this can improve total welfare because the more productive states are larger and have more low wage workers and unemployed, but a progressive government does not value transferring wealth from low to high income areas. Moving from centralization to the tiered system, there is an ambiguous effect on activity across states that depends on the relative progressivity of different governments.

To measure how states would be differentially affected under centralization compared to the current tiered system, the left three panels of Figure 2.7 show low-skilled employment, and high-skilled location decisions and wages change under these two regimes. Under centralization, Panel (a) shows that low-skilled employment is higher in most places. It mostly, but not exclusively, increases in states with low tiered minimum wages. The states experiencing the biggest decreases from centralization, represented by points far below the 45 degree line, are all relatively small, which reflects the government's trade off of redistributing to large workforces and low wage places. Without state minimum wages to reduce the high-skilled outmigration externality, the low-skilled workers in the lowest productivity areas are hit the hardest.

There is very little state heterogeneity in the impact on high-skilled wages because these workers are imperfectly mobile. If states became too different, some workers will move. Panel (c) shows that wages are close to the 45 degree line, but nearly uniformly above. With only a single low federal minimum wage of \$6.90, aggregate low-skilled employment increases in most states and drives up high-skilled wages. Centralization alone has only a small effect on the high-skilled migration externality and therefore allows for less redistribution from high to low-skilled workers, leading to higher average wages. However, since low-skilled employment decreases in some states, high-skilled workers' marginal products fall and they leave those affected low productivity areas. For this reason, Panels (a) and (e) look similar.

The regional distributional consequences of federal non-uniform over tiered policy setting are also ambiguous because states' preferences may be very different from the federal government. However, the optimal tiered minimum wages closely match the federal non-uniform, and so the right three panels of Figure 2.7 also show very little difference in low-skilled employment, and high-skilled location choices and wages.

2.6.5 A Brief Note on Decentralized Minimum Wage Setting

Competition between states for mobile high-skilled workers leads to low minimum wages under decentralized policy setting. Local governments are therefore less able to redistribute income from high to low-skilled workers. The amount of redistribution under decentralization compared with other systems importantly depends on regional heterogeneity in ν. Following the previous calibration for the federal government, we aim to estimate $\nu_j, j = \{1, ..., 48\}$, such that the observed 2015 minimum wage in state j maximizes j's welfare holding the other minimum wages fixed, so that the 2015 tiered policies are a Nash equilibrium. Since only 27 states set a binding minimum wage above \$7.25, ν_j is not point identified for every state.

In the calibrated model, there does not exist ν_j that replicates the state's 2015 policies. Under our assumptions and calibration, each state would prefer a smaller minimum wage, holding all other states' and the federal government's policies constant. That is, there does not exist a progressivity level such that the current state minimum wages are best responses. There are several possible reasons why the model is unable to match the data. First, governments may be more sophisticated. If a states believes that others will set lower minimum wages too when it lowers its own policy, then there are smaller gains from doing so. The main advantage of a lower minimum wage is to attract more high-skilled workers, but the amount is a function of low-skilled wages everywhere. Second, states could face political frictions. They may not be able to set their policies on a continuum but are instead restricted to a small grid. The observed policies may be optimal on the grid but not globally. Finally, the model is potentially misspecified or, at least, states may have different beliefs about the implied elasticities. In particular, the model implications for the effect of the minimum wage on low-skilled unemployment and high-skilled wages may not be true, for example if workers have monoposony power or the high-skilled have larger moving costs. More work is needed in the future to better understand state competition and minimum wage setting.

While we are unable to compare the decentralized and tiered equilibria explicitly, we aim to measure the relative trade-off of local and federal policy setting, holding preferences for redistribution fixed. Only 23 of the 48 states set binding minimum wages under decentralization when all governments have the same value of ν equal to 4.05. Without a binding policy in many states from the federal government, the high-skilled migration externality is large and states optimally set lower minimum wages. Figure 2.8 shows that the low-skilled wages in every state are lower under decentralized policy setting compared to the current US policies. However, total welfare under this version of decentralization is only 0.009 percent lower than the tiered system, or slightly worse than the centralized policy.

2.7 Conclusion

This paper provides a framework to think about the interaction of minimum wage setting and federalism. Previous optimal minimum wage research focuses on one jurisdiction, which misses the implications of government competition. We first build on the theoretical model of Lee and Saez (2012), extending it to a context in which high-skilled labor is mobile across regions. We present sufficient conditions for the desirability of a binding minimum wage, from the perspective of both a local and central government. Though minimum wages are not optimal in the presence of even a linear tax system in our model, their prevalence in the real world warrants consideration from an optimal policy perspective.

We calibrate a two-region model of a federation and compare minimum wage policies across four types of government structures: decentralization, centralization with uniform policy, a US-style combination of both, and a social planner. In our baseline model, the desirability of centralized policy setting is increasing in mobility; when mobility is shut down completely, decentralization obtains the social planners solution. However, when all movers face positive moving costs, and regions are heterogeneous, then the decentralized outcome leads to higher total welfare than federal uniform policy. Furthermore, suboptimal policy takes the form of minimum wages that are too low, such that the tiered system we observe in the US is weakly preferred to either decentralization or centralization exclusively. In the tiered system, the central uniform policy lessens the externalities from horizontal competition by local governments leading both types of governments to set policies more closely aligned with the social planner. Decentralized and centralized policy setting are strategic complements. When we extend our framework to US states, we find that the tiered system closely matches the social planner's solution. Our results indicate that higher levels of government should not forbid lower levels from setting their own minimum wages. These results are consistent across social welfare functions, although the effect of heterogeneity may depend on the progressivity of the social welfare function.

Though the US system performs well relative to other potential systems, its ability to efficiently redistribute income still falls short of a simple system of linear taxes. From the perspective of our calibrated model, the existence of minimum wages and redistributive taxes in the US remains a puzzle. Further research remains to be done on the types of models, economic or political in nature, which might give rise to the joint optimality of both types of policies. It is likely that the income tax is not set optimally due to political frictions, which lead to demand for other redistributive policies. For now, our results suggest the policy makers should pay attention to the mobility implications of minimum wage policies, as well as their interactions with the existing tax system.

Table 2.1: Calibration Results

Parameters	Estimate	Residual
α	.43	10^{-9}
$\bar{h}$	1	-
$\bar{\theta}$	.46	10^{-9}
ρ	0.29	-
$\bar{\xi}$	1.31	10^{-13}
ν	1	-

Table 2.2: Baseline Results

Authority	u	w_1^l	w_2^l	avg w^h	h_1	SW Loss
Panel A: $\gamma = 1$						
Competitive	.426	.361	.361	.657	.5	.23%
Decentralized	.426	.361	.361	.657	.5	.23%
Federal Uniform	.523	.370	.370	.643	.5	0
US	.523	.370	.370	.643	.5	0
Federal Non-Uniform	.523	.370	.370	.643	.5	-
Panel B: $\gamma = .995$						
Competitive	.429	.362	.358	.655	.506	.23%
Decentralized	.429	.362	.358	.655	.506	.23%
Federal Uniform	.533	.370	.370	.640	.509	.16%
US	.533	.370	.370	.640	.509	.16%
Federal Non-Uniform	.533	.372	.368	.640	.507	-
Panel C: $\gamma = .99$						
Competitive	.432	.364	.356	.653	.513	.23%
Decentralized	.432	.364	.356	.653	.513	.23%
Federal Uniform	.432	.364	.356	.653	.513	.23%
US	.432	.364	.356	.653	.513	.23%
Federal Non-Uniform	.522	.372	.364	.640	.513	-

Note: This table presents the simulated equilibria for our baseline calibrated model different minimum wage setting authorities and levels of regional heterogeneity. The competitive equilibrium is provided for comparison. Social Welfare (SW) Loss is defined relative to the Federal Non-Uniform, or Social Planner.

Table 2.3: $\gamma = .95$ and $\xi \sim \mathcal{U}[0.0025, \bar{\xi}]$

Authority	u	w_1^l	w_2^l	w^h avg	h_1	SW Loss
Competitive	.456	.374	.334	.638	.566	.21%
Decentralized	.497	.378	.338	.632	.566	.08%
Federal Uniform	.466	.374	.336	.367	.566	.19%
US	.542	.383	.343	.626	.566	0%
Federal Non-Uniform	.542	.383	.343	.626	.566	-

Note: This table presents the simulated equilibria for the case where $\gamma = .95$ and $\xi \sim \mathcal{U}[0.0025, \bar{\xi}]$ for different minimum wage setting authorities. The competitive equilibrium is provided for comparison. Social Welfare (SW) Loss is defined relative to the Federal Non-Uniform, or Social Planner.

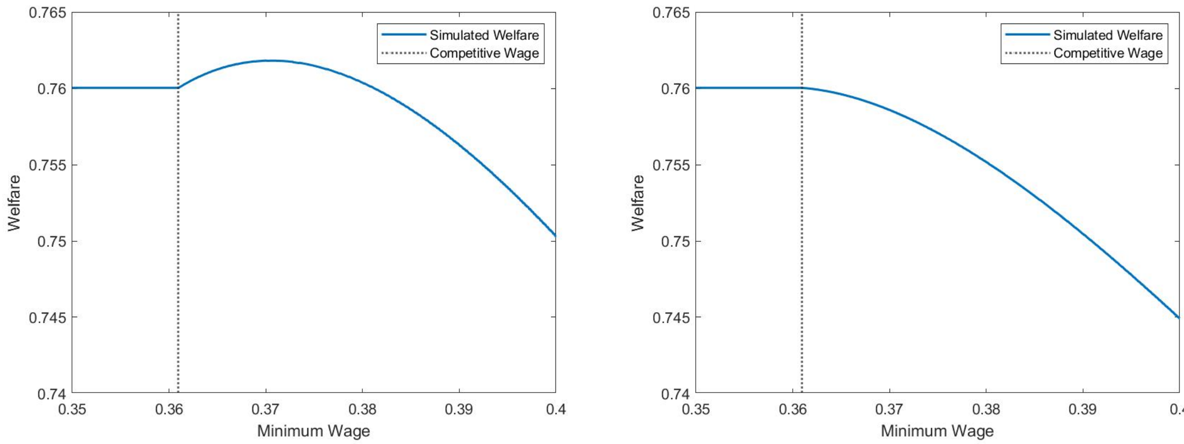
Table 2.4: High-skilled Migration Elasticity

	OLS $\log(p_j)$	First-Stage High-skilled Wage	IV $\log(p_j)$
High-skilled Wage	0.313***		0.954***
	(0.060)		(0.255)
Low-skilled Workers/100000		3.385***	
		(0.681)	
Observations	48	48	48

Note: The leftmost column (OLS) reports the estimated effect of average high-skilled wages in a state on the log of the percent of high-skilled workers in that state, using data from the 2015 American Community Survey. The middle column reports the first-stage for the IV regression in the rightmost column. Robust standard errors are reported in all columns.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5: US Model Calibration Results

National Parameters	Estimate	Residual
ρ	0.29	-
ν	4.05	0.012
State Parameters	Mean	Std. Dev.
$\bar{\theta}_j$	14.65	1.28
α_j	0.43	0.06
γ_j	21.44	2.72

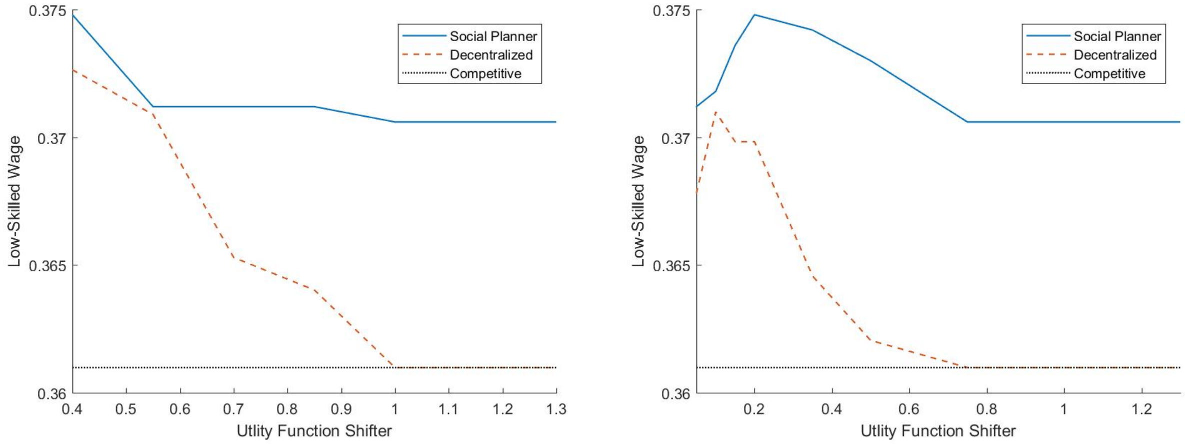


(a) Single Minimum Wage in Both Regions

(b) Minimum Wage in Region 1 Only

Figure 2.1: Welfare and the Minimum Wage

Note: Both panels of the figure plot the total social welfare in the economy for various levels of the minimum wage. The left panel plots welfare as a function of a single minimum wage applied in both regions while the minimum wage only applies in region 1 in the right panel. Since the regions are identical, the implications from a minimum wage in only 1 or 2 are the same. The vertical line denotes the competitive low-skilled wage in the absence of any policy.



(a) Baseline model

(b) Utility given by: $\log(2(c+\nu))$

Figure 2.2: Optimal Minimum Wage as a Function of ν

Note: In the right panel, we vary ν from 0.4 to 1.3 using the baseline calibrated model. In the left panel, we vary ν from 0.05 to 1.3 changing the social welfare function to $\log(2(c+\nu))$ The low-skilled wage in a jurisdiction is the maximum of the competitive wage and that government's minimum wage policy. The Social Planner's optimal policy (solid line) equals the Central Uniform and Tiered optimal policies since jurisdictions are symmetric.

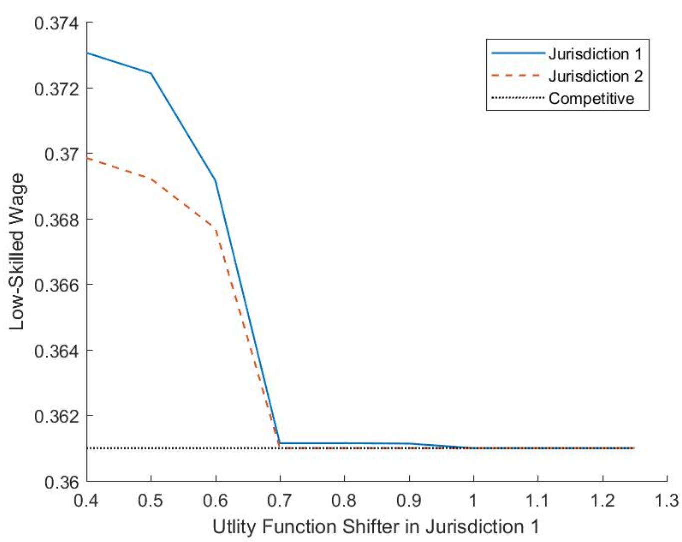
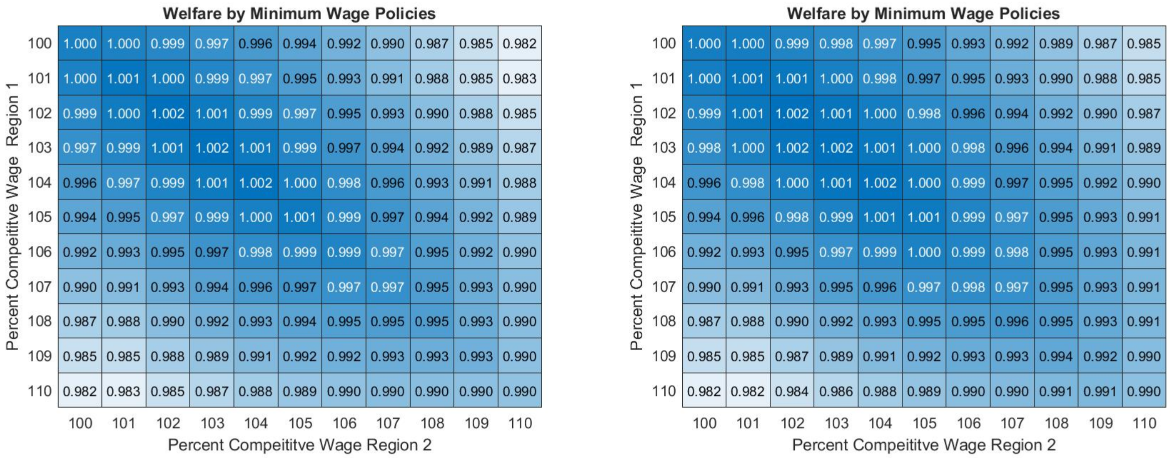


Figure 2.3: Optimal Minimum Wage as a Function of ν in Jurisdiction 1

Note: We vary ν in jurisdiction 1 from 0.4 to 1.3 keeping ν in jurisdiction 2 fixed at 1. The low-skilled wage in a jurisdiction is the maximum of the competitive wage and that government's minimum wage policy.

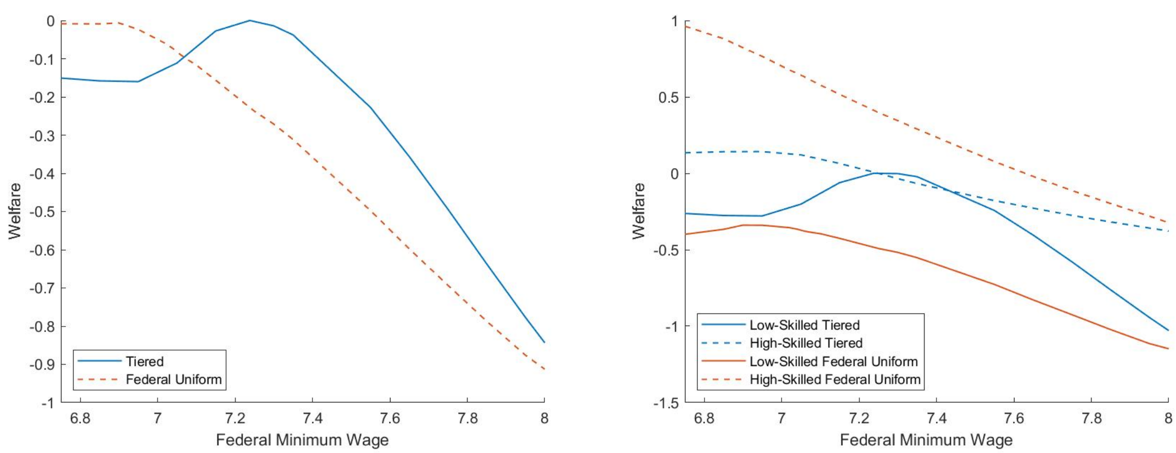


(a) $\gamma = 1$, $\min(\xi) = 0$

(b) $\gamma = .95$, $\min(\xi) = 0.0025$

Figure 2.4: Welfare and the Minimum Wage

Note: Panel (a) presents the total social welfare in the economy for different minimum wage policies in both regions ranging from the competitive wage to 110 percent of the competitive. Welfare is normalized so that it is 1 in the competitive case where neither region has a binding minimum wage. Panel (b) conducts the same simulations but for the case where $\gamma = .95$ and $\min(\xi) = 0.0025$ as in Table 2.3. Minimum wage combinations with higher total welfare are shaded darker.

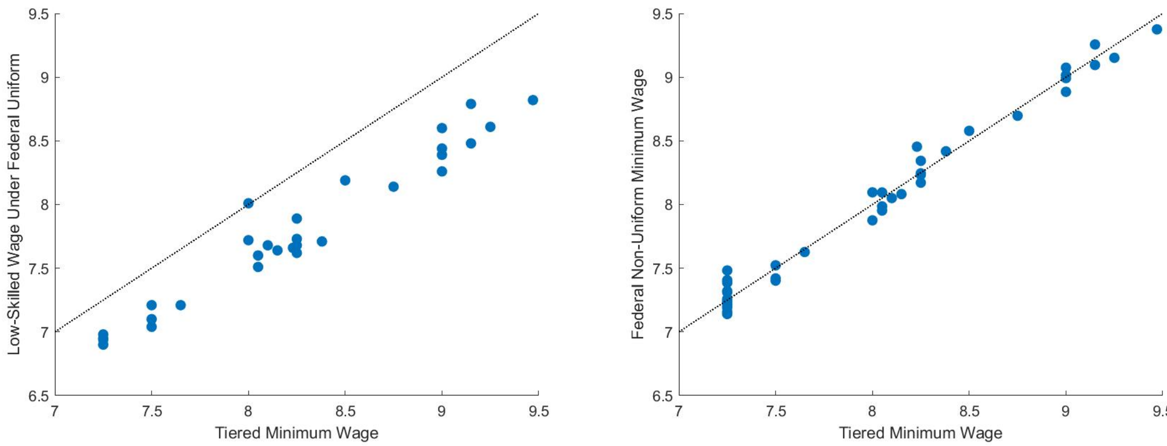


(a) Aggregate Welfare

(b) Welfare by Worker Type

Figure 2.5: US Welfare and the Federal Minimum Wage

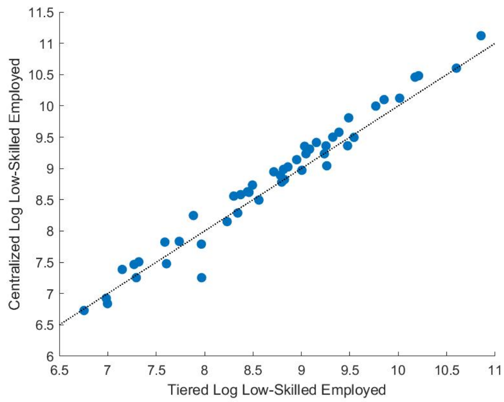
Note: Panel (a) plots total social welfare as a function of the federal minimum wage under tiered and federal uniform minimum wage setting authorities. Welfare is normalized to be 0 under optimal tiered policy setting with a federal minimum wage. Panel (b) splits welfare into the contributions from low and high-skilled workers separately. Welfare is normalized to be 0 for each worker type under optimal tiered policy setting.



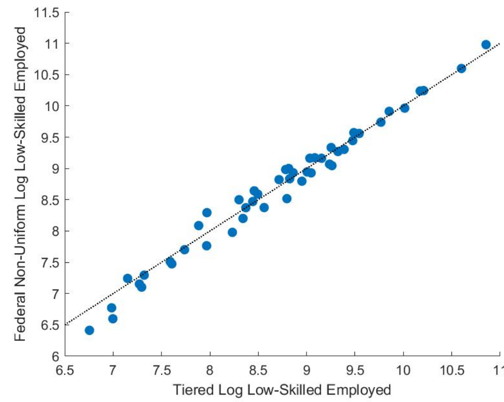
(a) Centralization (b) Federal Non-Uniform

Figure 2.6: Comparing US Minimum Wage Setting Authorities

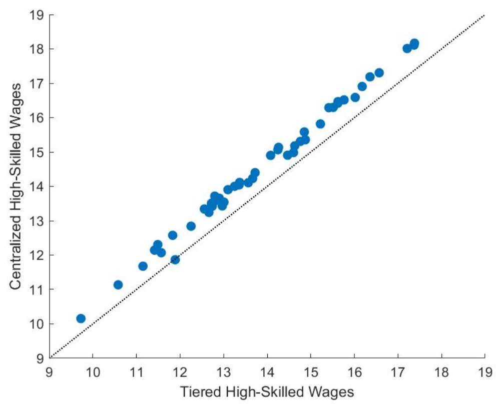
Note: Panel (a) plots the low-skilled wages in each state under tiered and federal uniform policy setting as well as a dotted 45 degree line for reference. Panel (b) similarly shows the relationship between tiered policy and the optimal federal non-uniform.



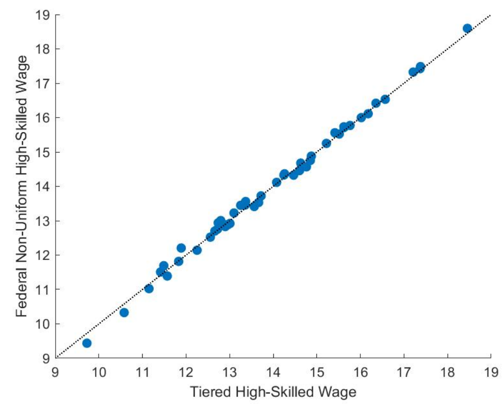
(a) Centralized Low-Skilled Employment



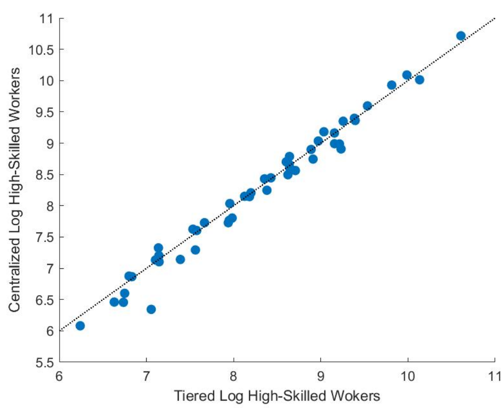
(b) Federal Non-Uniform Low-Skilled Employment



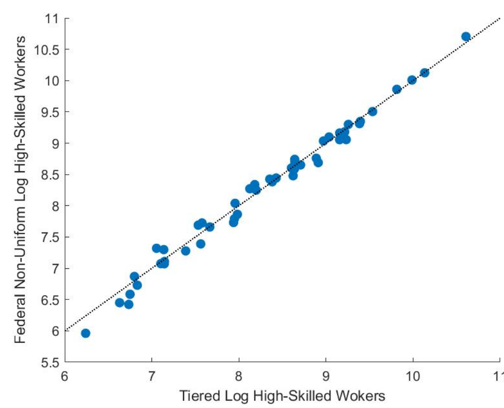
(c) Centralized High-Skilled Wages



(d) Federal Non-Uniform High-Skilled Wages



(e) Centralized High-Skilled Residencies



(f) Federal Non-Uniform High-Skilled Residencies

Figure 2.7: The Impacts of Minimum Wage Setting Authorities Across States

Note: The left three panels compare tiered and centralized policy setting, while the right three compare the tiered and federal non-uniform systems. All figures contain a a dotted 45 degree line for reference.

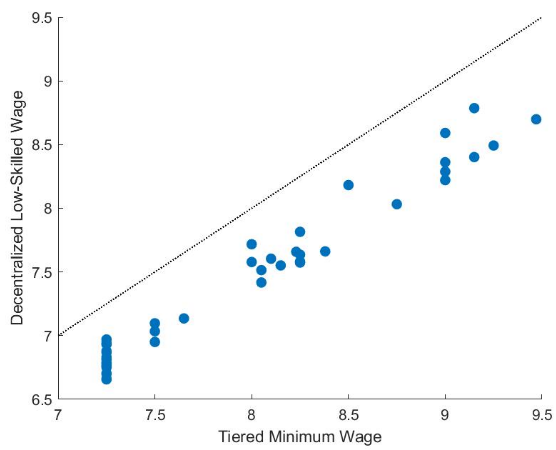


Figure 2.8: Decentralized Minimum Wage Setting

Chapter 3

College Choice, Private Options, and The Incidence of Public Investment in Higher Education (with John Bound)

3.1 Introduction

Understanding who benefits and who bears the burden of public investment in higher education is increasingly important for state and federal policy. Some states are greatly reducing support, while many Democratic 2020 presidential candidates called for tuition-free public college and an expanded Pell Grant. However, since the controversial findings of Hansen and Weisbrod (1969), the merits of public higher education as a redistributive tool have been called into question. They found that while higher education is mainly financed through progressive income taxes, take-up is predominately by higher income households. This result was supported by subsequent work in Wisconsin (Hansen 1970) and Florida (Windham 1970). Pechman (1970, 1972), using the same data as Hansen and Weisbrod, and more recently Bill Johnson (2006), using a nationally representative survey, argue that Hansen and Weisbrod misinterpret their findings. When stratifying by income, Pechman and Johnson show that net subsidies flow from high income families whose children often attend private colleges to the moderate income whose children often attend the publics.

The previous approaches to measuring incidence implicitly ignore general equilibrium considerations. Pechman and Johnson's method uses an accounting-style approach that sums the amount of money spent on an individual student from instructional expenditure and capital costs at public institutions and subtracts the amount paid in public college tuition and taxes. As Johnson acknowledges, these calculations leave out possible general equilibrium effects of state subsidies to public colleges. In particular, competition from public colleges affects who the private colleges admit and the tuition they charge. States' investments in public colleges have indirect effects on those who, in equilibrium, attend private colleges. Public investment in higher education also increases the size of the college educated workforce, affecting the relative wages of college and high school graduates.

How important are the general equilibrium implications of public investment for characterizing incidence? For what types of students do these effects matter the most? We build on the previous approaches to understand the redistributive implications of investing (or divesting) within the context of a general equilibrium model of higher education based on Epple et al. (2006, 2017, 2019) using a compensating variation (CV) framework. We find that incidence importantly depends on how the private market responds to changes in public funding, especially who private colleges admit and how they set their sticker prices, and on the effect that such changes have on relative wages. We calibrate the model to match national college attendance patterns and prices, which replicates the main findings from the accounting-style approach. This stylized setting allows us to study the role of college choice and market power for overall incidence in general equilibrium.[1]

Our framework addresses the previous concerns about the accounting approach by comparing the baseline college market equilibrium to one with lower levels of public investment in higher education. CV therefore measures how much additional income a household would need to be given, in excess of taxes, to be just as well off after a state spending decrease. Like the accounting approach, this method measures the direct effect of public investment on the value of enrolling in a public college. Our approach additionally captures the value of changes in price, quality, admissions decisions, and the changes in the labor market– the benefits are not only for those who attend.

We measure CV in two steps. First, we decrease state appropriations per student by 10

[1]We abstract from other relevant features like the importance of within and across state heterogeneity in the college market. We plan to explore the implications of cross border externalities in future work.

percent and hold private college sticker prices fixed. This exercise puts a lower bound on the extent to which private market power increases because the colleges can only raise tuition on a subset of students. However, with more aggregate demand, we expect the privates to increase their maximum prices as well. We then allow the sticker prices to adjust to highlight the importance of private market power for incidence. In doing so, we build on the previous models in the literature (e.g. Epple et al. (2006, 2017, 2019); Rothschild and White (1995); Chade et al. (2014), Fu (2014) and Fu et al. (2019)) by endogenizing sticker prices as a function of state funding for public colleges.[2]

While our results also imply that public investment is weakly progressive, we find that the accounting approach mischaracterizes incidence because the general equilibrium private college and labor market responses are large. Using our framework, we show that high-income-average-most-ability students benefit the most from public investment because they are only admitted to the highest quality private colleges when state spending is sufficiently high. As state spending decreases, the public universities charge higher tuition and become lower quality, which especially leads the low-income-high-ability students to switch sectors. The change in enrollment patterns from the spending cut crowds out the more modest-ability students. However, since they are very likely to attend the privates under the current spending level and pay high taxes, the accounting approach implies that they subsidize the system, rather than benefit.

Our results for the high-income-modest-ability students are driven by private college price setting behavior. The sticker prices imply that there is a minimum ability required for admission. These students are crowded out of the private market unless they are able to sufficiently compensate the colleges for their negative effects on peer quality.[3] This is only possible if the sticker price is high enough. As private market power and (sticker)

[2] Gordon and Hedlund (2018) also adapted this class of models to understand changes in tuition over time.

[3] In the model, college quality is partially determined by average student ability, or SAT scores. Students with ability above the college's average therefore create positive peer effects, while those below create negative.

prices increase, the minimum ability thresholds fall for two reasons. First, students can pay more in tuition to increase instructional expenditure and offset the negative peer effects. Additionally, at higher prices, fewer low-income-high-ability students enroll in the privates. Average peer quality does not increase as much and therefore neither do the negative peer externalities. Using policy variation in state investment over time, we find that private sticker prices only minimally increase when state support falls, and therefore these students bear much of the burden of divestment.

In aggregate, changes in these enrollment patterns affect the labor market returns to college as well. When state support decreases, the public colleges become more expensive and lower quality, while the privates restrict access through admission and prices. This makes college human capital scarce, driving up the college wage premium. When the privates increase their sticker prices in response to the decrease in public investment, even more students are priced out of the private market. This pushes up the college wage premium further. As college wages rise, so does private college market power and vice versa. Therefore, even students who never attend a public or private college benefit from the effects of public investment on the private market.

The increases in private market power and the college wage premium from decreases in state support have the largest benefit for high-income-high-ability students. These students are always admitted to the high quality privates and pay the full sticker price. When state support falls, their colleges become higher quality at a time when the labor market returns to quality are higher but pay nearly the same tuition and less in taxes. These students look very similar to high-income-modest-ability students based on the accounting approach, but the frameworks provide starkly different estimates of the welfare implications for these two groups. This is because the general equilibrium effects are large in magnitude, but the sign varies with income and ability. Appropriately accounting for these effects is necessary to measure incidence and the progressivity of public investment in higher education.

3.2 Measuring Incidence

Much of the previous empirical literature on the incidence of public investment in higher education (e.g. Bill Johnson 2006) uses an accounting approach. The net benefit captures a transfer to those who attend, measured by the sum of subsidies received from instructional and capital expenditure at public institutions less public tuition and tax payments. State appropriations directly affect this transfer value of public college attendance because it affects expenditure, tuition, and tax rates.

However, the accounting exercise misses important general equilibrium effects summarized by Figure 3.1. It only considers the farthest left arrow– a decrease in state appropriations creates a smaller transfer for those who attend public colleges. When the transfer value decreases, fewer students enroll. Instead, the lower in-kind subsidy leads some switch to the private sector (Peltzman 1973), while others do not enroll in any college. With increased relative demand, the private colleges have more market power– they are now able to charge tuition even further above marginal cost. A high quality, well-funded public puts competitive pressure on private colleges, which creates benefits for those students. When fewer students enroll after divestment, college-educated human capital becomes scarce and the college wage premium rises.

These mechanisms on college choice, private college market power, and the labor market returns to college importantly interact with each other, as indicated by the double sided arrows in Figure 3.1. For example, changes in enrollment patterns from state divestment affect college quality through peer effects, which further determines who the privates admit and how much they charge. Some students are crowded out of the private sector, further altering enrollment decisions. Implications in the college market then spill over to the labor market. Increased relative wages for college workers increase demand for higher education and therefore feed back onto private college market power. This compounds the effects on enrollment patterns and restarts the cycle.

This paper proposes a framework to incorporate the general equilibrium implications of state investment in public higher education to measure incidence. We first build on a stylized model of the higher education market from Epple et al. (2017). We then implement a compensating variation (CV) measure in the model that captures the amount of money one would need to give a household, in excess of tax and tuition revenue, so that it is indifferent between the current state spending per student levels and a 10 percent decrease. From the observed baseline setting, we simulate the general equilibrium effects of the divestment on the college and labor markets, as well as student enrollment choices. CV then computes the welfare effects by comparing these two equilibria.

Building on the accounting approach, the objective here is to measure the value inclusive of the general equilibrium effects, ν, in dollars, of higher education appropriations rather than just the direct transfer to those who attend. Let $U(y, b; z)$ denote the utility of student with parental income y and ability b when states spend z in appropriations per student. CV for counterfactual spending $0.9z$ compared to the baseline level is then

$$\begin{aligned} CV &= \nu - (t(z) - t(0.9z))y - (T(z) - T(0.9z)) \\ \text{such that } & U(y, b; z) = U(y + \nu, b; 0.9z) \end{aligned} \tag{3.1}$$

where $t(z)$ and $T(z)$ are taxes and tuition paid to public colleges as a function of the state spending, respectively.[4] Our analysis is similar to Epple et al. (2019) who measure the welfare effect of moving from a state with only the lowest quality public college to a state with four different levels of quality. However, their analysis does not capture the general equilibrium effects on the college and labor markets that are the focus of our analysis.[5]

[4] In our model presented in the next section, all state income taxes support the public colleges, but only the fraction of federal income taxes used to finance financial aid at the public colleges does. Tuition paid equals the sticker price minus federal financial aid received.

[5] Kasman and Guyot (2019) also investigate the implications of college subsidies using an agent-based model, although their main interest is enrollment and their framework does not consider private college market power.

3.3 The Higher Education Market

We build on the general equilibrium model of higher education from Epple et al. (2017) with two main extensions that highlight the importance of private college market power and the labor market returns to college to measure incidence. First, we endogenize sticker prices. In doing so, private colleges are better able to adjust their prices and admissions decisions with changes in government policy to further increase their market power. Second, we model the value of college as a function of the size of the college educated workforce. Changes in aggregate enrollment affect an individual's demand for higher education through its impact on the college wage premium.[6] We additionally specify a more flexible parameterization of the model that is better able to match attendance by income and ability.

The market consists of three private colleges, one public college in each of two states, a unit mass of households, and state and federal governments. We describe each in turn.

3.3.1 Households

There is a unit mass of households living in two identical states. Households are characterized by their state of residence, s, parental income net of college expenses[7], y_p, and student ability, b, which are drawn from a joint distribution $f(b, y_p)$. Students make college enrollment decisions based on the wage adjusted value of human capital y_{sj} from enrolling in j with quality q_j and the cost. College j costs the household $p_{sj}(b, y)$ in tuition, and L in living expenses. Federal financial aid of A_{sj} is available to offset college costs. Households attend

[6] George Johnson (1984) shows that this general equilibrium effect can be so strong that non-college enrollees prefer more government investment in public colleges than those who directly benefit.

[7] All dollar amounts are measured in tens of thousands.

the option that maximizes their utility:

$$U_j(s,b,y,\epsilon_j) = \ln(y_{pj}^{\alpha_p} y_{sj}^{\alpha_s}) + \epsilon_j \tag{3.2}$$

$$\text{where} \quad y_{pj} = (1 - t_s - t_f)y - p_{sj} - L_j + A_{sj} \tag{3.3}$$

$$\text{and} \quad y_{sj} = \begin{cases} w_c \beta_1^c \big(\beta_2^c b^{\beta_3^c} + (1-\beta_2^c) q_j^{\beta_3^c}\big)^{1/\beta_3^c} & \text{if attends college j} \\ w_h \beta_1^h b^{\beta_2^h} y^{\beta_3^h} & \text{if does not attend college} \end{cases} \tag{3.4}$$

where α_p and α_s are weighting parameters that reflect the relative values of college costs and future student earnings, t_s, and t_f are state and federal taxes, w is the wage value of college or non-college human capital, and the disturbances ϵ_j are i.i.d. Type I Extreme Value preference shocks for college j. The value of attending college depends on a student's ability and college quality, while the value of the outside option depends on ability and parental income. Under standard regularity conditions on ϵ_j, as in McFadden (1974), we can recover the probability that a student of type (b, y) attends college j, r_{sj}, given the set of admitted colleges J and the outside option of not attending college O.[8]

The labor market value of college human capital, w_c, and non-college human capital, w_h, are determined by a CES production function in aggregate college and non-college educated labor. The college wage premium for a unit of human capital is:

$$\frac{w_c}{w_h} = \frac{s}{(1-s)} \left(\frac{C}{H}\right)^{\rho - 1} \tag{3.5}$$

where s is the share on college human capital, ρ determines the elasticity of substitution, C and H are the aggregate amounts of college and high school human capital in the economy, respectively.

[8] $r_{sj} = \frac{y_{pj}^{\alpha_p} y_{sj}^{\alpha_s}}{\sum_{k \in J \cup O} y_{pk}^{\alpha_p} y_{sk}^{\alpha_s}}$. If a student does not attend college, L, A and p are 0.

3.3.2 Private Colleges

The three private colleges admit students and choose individual prices to maximize quality, q_j, subject to a budget constraint and a maximum or sticker price P_j^c. College quality is a function of mean student ability, θ_j, and instructional expenditure, I_j.

$$q_j = \phi_j \theta_j^{\gamma} I_j^{\omega}, \gamma, \omega > 0 \tag{3.6}$$

Mean student ability and instructional expenditures per student are important inputs to the education production function and common proxies for quality (Black and Smith 2006). The colleges' objective function is similar to prestige maximization, where each wants to enroll high test score students that raise its national profile, while also collecting as much revenue as possible that can be used to improve instructional quality or as merit aid to attract more high test score students.

All colleges, public and private, have the same cost function,

$$C(k, I) = F + v_1 k + v_2 k^2 + kI \tag{3.7}$$

where F are fixed costs, k is the number of enrolled students, and v are parameters to be calibrated that capture the marginal cost of enrolling more students. Private college j has exogenous endowment E_j to spend on instruction in addition to collected tuition. A student with ability b and income y faces price p_{sj} that depends on her effective marginal cost,

$$EMC_j = v_1 + 2v_2 k + I_j + \frac{\partial q_j/\partial\theta}{\partial q_j/\partial I}(\theta_j - b) \tag{3.8}$$

and the conditional choice probability r_{sj} that a student of her type (b, y) attends college j

$$p_{sj}(b, y) + \frac{r_{sj}(b, y)}{\partial r_{sj}(b, y)/\partial p_{sj}(b, y)} = EMC_j(b) \tag{3.9}$$

but must not exceed sticker price $P_j^c(z_s)$, which is a function of state appropriations, z_s. EMC depends on ability through peer effects. Private colleges want to charge high ability students lower prices so that they are more likely to attend and raise θ. A student is admitted if her marginal revenue (the minimum of the left hand side of (3.9) and the sticker price) is greater than effective marginal cost. Colleges must satisfy their budget constraint. They spend tuition and endowment income to cover custodial and instructional costs.

3.3.3 Public Colleges

Each state operates a public college. The public college in state s maximizes expected future income of its state's residents, y_{sj}, subject to a budget constraint.

Public colleges also take appropriations, z_s, to be exogenous. Given z_s, the public college in state s charges T_s to residents and T_{so} to non-residents. In order to raise sufficient funds for instruction, colleges adjust tuition based on changes to z_s. Colleges may also change the relative proportion of residents and non-residents admitted. We assume tuition is a function of appropriations such that

$$\begin{aligned} T_s &= \left[1 + \kappa_s\left(\frac{z_s - \overline{z}_s}{\overline{z}_s}\right)\right]\overline{T}_s \\ T_{so} &= \left[1 + \kappa_{so}\left(\frac{z_s - \overline{z}_s}{\overline{z}_s}\right)\right]\overline{T}_{so} \end{aligned} \tag{3.10}$$

where κ is the elasticity of tuition with respect to appropriations, and $\overline{T}$ represents baseline equilibrium tuition level. This equation is only locally valid and is therefore only used to understand tuition setting under small changes in appropriations. All students of the same residency type at a public college face the same sticker price, although students with different incomes pay different amounts due to differences in federal financial aid. Tuition and appropriations per student fund instructional expenditure per student and custodial costs.

Epple at al. (2017) show that the public college j admits all students with ability above a residency specific threshold for residents, b^s_{min}, and non-residents, b^o_{min} respectively implicitly defined by:

$$y_{sj}(q_j, b^s_{min})/\lambda + T_s + z_s - EMC_j(b^s_{min}) = 0 \tag{3.11}$$

$$T_{so} + z_s - EMC_j(b^o_{min}) = 0 \tag{3.12}$$

where λ is the multiplier on the budget constraint. The non-resident threshold is a function of tuition and the appropriations level, while the resident threshold also accounts for how educating student with ability b changes the total stock of in-state human capital.

3.3.4 Federal and State Governments

The federal government distributes financial aid, A_{sj} through a Pell Grant type program, that takes into account expected family contribution as a function of income $EFC(y)$. A student can receive at most $\bar{A}$ in aid.

$$A_{sj} = \min\{\max[0, P^c_j + L - EFC(y)], \bar{A}\}$$

The state and federal governments must balance their budgets. Let Y_s be pre-tax aggregate income in state s.

$$t_s Y_s = z_s k_s, \quad s = \{1, 2\}$$
$$t_f \sum_s Y_s = \sum_s \int \sum_j r_{sj} A_{sj} f_s(b, y) db dy$$

where k_s is the number of students who attend the public college in state s, regardless of each student's state of residence. State colleges can give admission and tuition preference to residents but do not treat students differently after enrolling.

3.3.5 Equilibrium Definition

An equilibrium consists of a price and quality vector with corresponding college characteristics (θ_j, I_j, k_j) for all $j \in J \setminus O$; state admission criteria for all states for residents and non-residents; state of residence s for all households, which gives $f_s(b, y)$; and a set of student choices for all (s, b, y) and $j \in J$ and choice probabilities $r_{sj}(b, y)$ such that:

1. Private colleges choose who to admit and what prices to charge to maximize quality, taking the prices and qualities of the other colleges, the student choice probability functions, and public policies as given.

2. Public colleges choose who to admit to maximizes in-state achievement, taking the prices and qualities of the other colleges, the student characteristics, and public policies as given.

3. Household (b, y) pick where to attend to maximize utility, taking college prices and qualities, public policies, and the decisions of other households as given.

4. State and federal budgets are balanced.

3.3.6 Brief Discussion of Model Limitations

The model outlined in this section is highly stylized, but the general equilibrium structure allows us shed light on the importance of private college market power and the labor market for measuring the incidence of public investment in higher education, which is missing from the previous measures. Before discussing how the model is calibrated and the results, it is important to understand some of the model's key simplifications.

First, the public sector only consists of one college per state while states have several colleges of varying quality. Including additional public colleges would provide more competitive pressure on the private sector and therefore strengthen the mechanisms we highlight. Similarly, public colleges do not offer institutional aid in the model, which also understates the competitive pressure they provide. Cook and Turner (2021) show that public research universities have progressive pricing policies, although that is not the case for non-research universities. If the public universities could attract more high-ability students away from the privates with subsidies, then the privates would have less market power. However, the opposite is true for the size of private sector– including more privates would increase the competitiveness of market and therefore lower the value of public competition.

Related to our assumptions on the size of the market, we also assume colleges do not face capacity constraints. With capacity constraints, public colleges have less resources to spend on instruction but also no longer enroll the lower ability students who lower peer effects but whose education raises average state achievement. This leads to an ambiguous quality change and therefore an unknown change in public value and competitive pressure. Private college capacity constraints also have an ambiguous effect on market power because it restricts access but also lowers quality and therefore decreases both demand and supply.[9] Most importantly for incidence, when private college demand increases after a state funding cut, capacity constraints would lead fewer total students to enroll. Excluding this constraint would lead to larger negative welfare effects on the baseline ($z = \bar{z}$) marginal ability students who have high effective marginal costs.

Our simplifications on the household side of the market are also important for incidence. Households are only able to pay for college out of current income and so some low-income households may be financially rationed out of the private market, especially when the returns to college are high. We may therefore overstate the value of public colleges to the modest-income-high-ability students who are always admitted to the privates but can only afford to attend in our model with public competitive pressure.[10] However, a loan market weakly increases households' willingness-to-pay for college and therefore also increases private college prices and market power. Overall, we find that changes in private market power from decreases in public investment mainly affect high-income-modest-ability students who are always able to afford college without loans. A loan market would therefore give greater college access to low-income-higher-ability students, exacerbating the main mechanisms we study in this paper.

[9] Since the privates maximize quality in the model, a constraint necessarily would weakly lower quality.

[10] Our framework does not capture the impact of liquidity constraints and household borrowing for college, which are likely important (see Lochner and Monge-Naranjo 2012) and vary across the income distribution.

3.4 Private Colleges and Incidence

Public investment in higher education creates general equilibrium implications for the private college market. To measure the role private colleges play for measuring incidence, first consider how appropriations, z_s, affect their prices. For the purposes of this illustration, we hold the labor market fixed. Given the model's parametrization and Equation 3.9,

$$p_{sj} = \frac{(1-r_{sj})\alpha_p}{1+(1-r_{sj})\alpha_p} EMC_j(b) + \frac{y - L_j + A_{sj}}{1+(1-r_{sj})\alpha_p} \tag{3.13}$$

and therefore

$$\frac{\partial p_{sj}}{\partial z_s} = \underbrace{\frac{\alpha_p(y - L_j + A_{sj} - EMC_j(b))}{(1+(1-r_{sj})\alpha_p)^2}\left(\frac{\partial r_{sj}}{\partial z_s}\right)}_{\text{change in relative value of private college}} + \underbrace{\frac{(1-r_{sj})\alpha_p}{1+(1-r_{sj})\alpha_p}\left(\frac{\partial EMC_j(b)}{\partial z_s}\right)}_{\text{change in cost to college of enrolling}} \tag{3.14}$$

where

$$\frac{\partial EMC_j(b)}{\partial z_s} = \underbrace{2\nu_2 \frac{\partial k_j}{\partial z_s}}_{\text{change in custodial costs}} + \underbrace{\left(1 + \frac{\gamma}{\omega}\left(1 - \frac{b}{\theta_j}\right)\right)\frac{\partial I_j}{\partial z_s}}_{\text{change in expenditure costs}} + \underbrace{\frac{\gamma I_j b}{\omega \theta_j^2}\frac{\partial \theta_j}{\partial z_s}}_{\text{change in peer externality}} \tag{3.15}$$

and $\partial r_{sj}/\partial z_s$ is determined by the household's utility function.[11] When appropriations fall, instructional expenditure at the public colleges fall as well and tuition increases. Households are more likely to enroll in the privates as a result. This implies $\partial r_{sj}/\partial z_s < 0$ and $\partial k_j/\partial z_s < 0$ as well. The first term in Equation 3.14 captures the first order effect of changes in the value of a student's other college options on private college prices. When the value of the public universities decreases and they provide less competitive pressure, the privates gain market power and increase prices.

The market level changes also create general equilibrium effects from student sorting, reflected in the second term of Equation 3.14 and broken down into three components in

[11] The derivate is only defined for students who attend college j in the baseline. For those students, $y - L + A_{sj} - EMC_j(b)$ is positive since price is less than or equal to $EMC_j(b)$ and students only attend if they can afford to.

Equation 3.15. First, decreases in public college appropriations increase custodial or operating costs as long as the college has increasing marginal costs, so that v_2 is positive. The middle term reflects changes in instructional costs. If decreases in appropriations allow the college to charge higher prices because $\partial r_{sj}/\partial z_s < 0$, it will have more revenue to increase instructional expenditure. The direct financial cost of enrolling students will increase. The impact of changes in instructional expenditure on prices is weighted by a function of student ability since instruction and average ability are linked through the quality production function. Since expenditure and average ability are complements, the value to the college of enrolling high-ability students or cost of enrolling low-ability students is larger when expenditure is higher.

The final term in Equation 3.15 captures how general equilibrium sorting of students affects relative college quality from changes in peer effects. The sign of $\partial \theta_j/\partial z_s$ and the final term therefore depends on who the marginal students to private colleges are. When the public colleges become lower quality, low-income-high-ability students may then find it worthwhile to pay for the more expensive privates. In that case, θ_j will increase. The opposite is true if high-income-modest-ability students are on the margin. When θ_j is larger, $\theta_j - b$ is also larger and low-ability students produce more negative peer externalities. They must pay higher prices to compensate, although colleges are unable to raise tuition above a maximum P_j^c. Overall, this discussion implies that the sign of $\partial p_{sj}/\partial z_s$ is theoretically ambiguous and potentially varies across the income and ability distribution. However, in our calibrated model presented in the next section, we find that both $\partial p_{sj}/\partial z_s$ and $\partial EMC_j(b)/\partial z_s$ are negative for all students.

State investment in public higher education additionally affects who the privates admit. Each private admits a student if her effective marginal cost is less than the minimum of the optimal price given by Equation 3.9 and the sticker price, that is if $\min(P_j^c, p_{sj}(b, y)) \geq EMC_j(b)$. Students that pay more than their effective marginal cost but less than the sticker price before a small appropriations decrease will still be admitted afterwards. The spending shock has an ambiguous effect on welfare for these students since they now face higher prices, but also benefit from higher quality private colleges.

The students most impacted by decreases in state spending are those who are no longer admitted to the privates. The sticker price and effective marginal cost together imply a minimum ability threshold at each college defined by $EMC_j(b) = P_j^c$. Who is still admitted after the state spending change depends on how both pieces respond. Holding the sticker prices fixed, a decrease in appropriations will increase the minimum ability threshold if $\partial EMC_j(b)/\partial z_s < 0$ for the marginal ability student. Students who can no longer attend their best option face a potentially large welfare loss. How many students are affected depends on the private college sticker price responses. If the privates also raise their maximum prices when appropriations fall, $\partial P_j^c/\partial z_s < 0$, then more low-ability students can compensate the privates for their high effective marginal costs. In the remainder of the paper, we calibrate the model and present counterfactual simulations to understand the importance of these mechanisms for incidence.

3.5 Calibration and Equilibrium Description

We calibrate the model to match the college going behavior across the income distribution and previous estimates from the literature that describe how students and institutions respond to different higher education policies. The estimates are presented in Table 3.1. First, using information on the joint income-ability distribution from the High School Longitudinal Study of 2009 (HSLS), we estimate the income distribution and the correlation between income and ability, assuming that they are jointly log-normally distributed. Following Epple et al. (2017), we assume that the ability distribution has a mean of 1 and a standard deviation of 0.15.[12] We truncate the income distribution at the lowest bin we observe in the data ($15K) so that all households can afford the in-state college option as well as $300K.

Many baseline parameters, specifically government policies, sticker prices, and endow-

[12]Using the PowerStat tool from the National Center for Education Statistics, data is available by the following income bins: less than $15K; in increments of $20K up to $235K; more than $235K. We measure ability using the mathematics quintile score. Income is scaled to 2013-2014 using the Consumer Price Index.

ments, are directly observed in the data. The calibration uses the 2013-2014 academic year to align with the HSLS. The Delta Cost Project has data on institution-level resident and non-resident tuition, state appropriations, and private college sticker prices. We compute the enrollment-weighted average for public four-year colleges and private not-for-profit four-year colleges, divided into three groups by SAT scores. Living Expenses are from the National Center for Education Statistics (NCES). Federal financial aid policies and private college endowments follow Epple et al. (2017)[13].

We calibrate the remaining household and college parameters, $\alpha s, w\beta s, \phi s, \gamma, \omega, F$, and vs, to minimize the sum of squared percent deviations of: (1) average private college tuition of \$27220 (National Postsecondary Student Aid Study 2012, Epple et al. 2019 scaled to 2013-2014 dollars); (2) private school enrollment as percent of total enrollment of 28.5 (HSLS); (3) enrollment elasticity of institutional aid of 0.7 (Van der Klaauw 2002); (4-5) estimates of the elasticity of aid with respect to ability and income from Epple et al. (2019); (6) the 20 percent earnings increase for the marginal-ability student from attending a public college (Hoesktra 2009 and Zimmerman 2014)[14]; (7) public college instructional expenditure per student of \$10463 (Delta Cost Project); (8-13) the percent of students from each of 6 income bins who enroll in any 4 year college weighted by the percent of students in each bin[15]; and (14-18) the percent of students from each ability quintile who enroll in any 4 year college weighted by the percent of students. We additionally set v_1 to 0, based on previous results from Epple et al. (2006, 2017). Appendix C.1 presents the model fit.

From the calibrated equilibrium, the returns to college quality and ability, βs, and the labor market returns to human capital, ws, are not separately identified. While βs are structural parameters, ws are equilibrium parameters and functions of public investment in

[13] The maximum federal aid is set to \$4K to match estimates from the 2011-12 National Postsecondary Student Aid Study for those who applied for federal aid.

[14] The model implies a minimum-ability threshold at public colleges. This creates a regression-discontinuity, where students just above the threshold have access to the public colleges, but those below do not. We match the model implied earnings gain from college attendance for these students, based on β, to the RD estimates.

[15] We combine the highest income bins to create \$95K to \$135K and more than \$135K so that all bins have at least 10 percent of the population. We exclude the lowest income bin of less than \$15K since we truncate the income distribution at that point.

higher education. Assuming that the labor market value of human capital is determined by a CES production function in college and non-college educated labor, we set ρ to 1.7 (Autor 2014). We then recover the share parameter, s, assuming that the entire baseline earnings difference between workers is because of human capital differences or that $w_c/w_h = 1$.[16] Under these assumptions, we recover βs, w_c and w_h for the baseline equilibrium.

Finally, we calibrate how public and private college tuition respond to changes in public investment in higher education. We set $\kappa_s = -0.263$ and $\kappa_{so} = -0.066$ based on Bound et al. (2019).[17] In the next subsection, we estimate the private college sticker price elasticity with respect to state appropriations, κ_p.

3.5.1 Sticker Price Response to Public Investment

The model implied incidence of public investment in higher education depends on the responsiveness of private college sticker prices to changes in state funding. The progressivity of public investment is a function of the magnitude of the elasticity since a larger price response implies a smaller change in the minimum ability thresholds at the private colleges. We estimate the elasticity of private college sticker prices with respect to state appropriations, κ_p, from the following specification:

$$\log(P^c_{j(s)t}) = \kappa_p \log(z_{j(s)t}) + \pi_{j(s)} + \psi_{st} + u_{jt} \tag{3.16}$$

$$\text{where } z_{j(s)t} = \sum_i z_{it} \frac{k_{ij,1998}}{k_{j,1998}} \tag{3.17}$$

where $P^c_{j(s)t}$ is the sticker price of private college j in state s in year t, z_{it} is the average appropriations in state i and year t, $z_{j(s)t}$ computes the relevant state appropriations for college j by weighting z_{it} by the fraction of j's students from i in the baseline year 1998, and

[16] From the model's calibrated equilibrium, we cannot recover both s and the college wage (per unit of human capital) gap. However, since college educated labor is a scarce resource, we expect $w_c \geq w_h$. From calibration moment (6), the marginal student's earnings gains also imply that $w_c \leq 1.2 w_h$ if college weakly increases human capital. We therefore choose $w_c/w_h = 1$ as a normalization. For a small policy change, like in our counterfactuals to measure incidence, the equilibrium wage ratio is mainly determined by ρ.

[17] Rizzo and Ehrenberg (2004) find that for a \$1000 increase in appropriations, resident tuition decreases by \$60, while non-resident decreases by about \$222. Webber (2017) finds larger elasticities, especially for more recent years, but does not differentiate between resident and non-resident tuition.

$\pi_{j(s)}$ and ψ_{st} are institution and state-by-year fixed effects, respectively.

Equation 3.16 is similar to a shift-share design where z_{it} captures yearly variation in the effective treatment for college j and $k_{ij,1998}/k_{j,1998}$ captures j's dependence on students from i. As appropriations and public college quality in i decrease, the demand for private options increases, allowing those colleges to raise prices. The importance of a spending shock in i to college j depends on how many students from that state tend to enroll.[18] This strategy exploits within state variation based on where private colleges draw their student populations.

McDuff (2006) provides empirical support for these mechanism where students respond to changes in public investment in higher education. He finds that a 1 standard deviation increase in public college quality, which is measured as a combination of mean SAT scores, freshman retention, and spending per student at the flagship, increases the percentage of scores sent to in-state public schools by about 6 percentage points and the likelihood of enrolling in an in-state public school by 5 percentage points. Increased demand in one sector implies relative decreases in others, the private market or no college. McDuff also finds that lower income households are more price sensitive, while the higher income are more quality sensitive. This evidence is consistent with our calibrated model and we highlight its implications for enrollment and incidence in the next section.[19]

We estimate Equation 3.16 using a balanced panel of Bachelor's, Master's and Doctoral private colleges from 2000 to 2017. For each institution, the Integrated Postsecondary Education Data System (IPEDS) reports the college's sticker price in a given year, as well as the number of students from each US state who attend. Grapevine, an annual compilation of data on state fiscal support for higher education, from the Center for the Study of Education Policy at Illinois State University and the State Higher Education Executive Officers

[18]Contemporaneous measures of the share of students from each state would lead to an upward bias since enrollments are also a function of the sticker prices. We therefore use 1998 as a base year to calculate the weights.

[19]Bound et al. (2009) provide additional suggestive evidence that students responded to the decrease in public college quality over time.

(SHEEO) report the state-level average appropriations per student at public colleges. We allow κ_p to vary by quartile of SAT scores since our model predicts that the elasticity depends on how close substitutes the public and privates colleges are.[20] The results are reported in Table 3.2. Our preferred specification is column (3) with state-by-year fixed effects. We present specifications with only year fixed effects, and both year fixed effects and state linear time trends to show the importance of capturing business cycle fluctuations. When the economy is doing well, states have more resources to spend on appropriations, but so do households to spend on education, which would lead κ_p to be biased upwards.

Our results suggest that private colleges adjust their maximum prices to further increase market power when state appropriations fall, but the effect is small in magnitude. A 10 percent decrease in appropriations would lead to less than a 1.3 percent increase in private college sticker prices. The effect also appears to be concentrated in the upper-middle quartile of selectivity, at institutions like Northeastern University. Colleges at this level of selectivity are likely close substitutes for the publics. The student-weighted average SAT scores at public colleges is at the boundary of the lower-middle and upper-middle quartiles and well within the upper-middle for flagships. Our results may understate the effects at less selective institutions since many draw students locally, which makes it difficult to detect a response in our framework with state-by-year fixed effects. Some of these institutions are also horizontally differentiated, religious colleges for example, and therefore may also not directly compete with publics. Highly selective institutions compete less with the public sector and respond less as well, based on our estimates.

Based on the results, we set $\kappa_p = -0.126$ for all private colleges in the model. This subtle response may be consistent with quality-maximizing private colleges if high-ability-low-income students' application decisions are sensitive to the sticker price, as has been well-documented (e.g. see Bettinger et al. (2012), Hoxby and Avery (2013), Hoxby and Turner (2015), and Dynarski et al. (2018)). Some private colleges have recently lowered

[20] IPEDS reports the 25th and 75th percentiles of the math and verbal section of the SAT. We measure the institution's SAT scores as the average of all four components in 2001, the first year available in the data.

their sticker prices in hopes of attracting more students (Anderson 2019). Colleges may also have a different objective function for setting their maximum prices. Epple et al. (2006) suggest that lower than quality-maximizing prices would reduce the number of low-ability students who harm the college's reputation, limit the private incentives of administrators to admit students of wealthy households for personal gain, and lead to higher alumni giving.

Our results are also consistent with previous evidence on average prices, as in Fillmore (2016). Using the National Postsecondary Student Aid Study, he finds that, all else equal, colleges offer lower prices to students when they believes they face stiffer competition. Every additional school listed on a student's FAFSA is associated with a \$373 increase in institutional aid. This effect is primarily driven by private colleges and the estimated association for very selective privates is more than \$900 per additional school listed. Large changes in the market that affect every student's application portfolio should then have general equilibrium consequences for the entire tuition distribution, pushing up sticker prices as well.

3.5.2 State Investment and The Higher Education Market

State investment in public higher education has important implications for the entire college market and the labor market. Using the baseline calibrated model with state spending $\bar{z}$ and private sticker prices, P_j^c, we simulate the effects of a 10 percent decrease in public support with and without endogenous sticker prices. Table 3.3 includes key features of the equilibria to highlight the general equilibrium consequences of state divestment. The next section then quantifies the magnitude of these different impacts across the income and ability distribution to understand who benefits and who bears the burden of public investment.

Based on our calibration (column 1), the market consists of an open-access public college in each state, two nearly open-access private colleges, and one selective private college. We refer to the privates as LP, MP, and HP for low-quality, middle-quality, and high-quality private colleges. When state appropriations fall to $0.9\bar{z}$ (column 2), the public colleges charge higher tuition to recover lost revenue, but spend less on instruction. These two changes depress aggregate public college demand and have two implications for private market. First, the privates charge higher prices to increase instructional expenditure and quality, further driving up their own demand for most. However, prices increase too much for the low-income, leading some to substitute to the in-state public, raising its average student ability.[21] The general equilibrium effects of state investment on the private market therefore feedback onto the public colleges.

Second, some baseline public college students now enroll in the privates. In particular, modest-income-high-ability students, who receive substantial private institutional aid, switch sectors. The change in enrollments increases the average student ability at the highest quality private, which raises the minimum ability threshold. Those between the 60th and 62nd percentiles of the ability distribution can no longer sufficiently compensate the HP college at the baseline sticker prices, and are crowded out.

Increased selectivity at the highest-quality private college creates negative spillovers on the other privates. The modest-ability students, or those just above the 60th percentile, no longer admitted after the state funding decrease are now likely to enroll in one of the other privates, lowering their average student ability. This effect outweighs the increased demand from the modest-income-high-ability students who attend the public colleges when state support is high. Overall quality at 3 and 4 still increases because the colleges enroll more high-income students and can charge higher prices from the change in demand.

Based on the implications for student sorting and college quality, state divestment in public higher education decreases total college human capital accumulation. The implied college wage gap increase by 0.4 percent induces some baseline non-college workers to enroll. Even though enrollment goes up, most attend the now lower quality publics. Although more students would want to enroll in the private sector, since the returns to any college and college quality increased, the privates limit access through their admissions decisions and prices.

When private colleges increase their sticker prices slightly after the decrease in state

[21] The shift of more lower income households to the the less expensive public sector or to no college also decreases total federal financial aid.

funding (column 3), more modest-ability students can sufficiently compensate the highest quality private for their effect on θ_{HP}, and are now admitted. The minimum ability threshold falls slightly but is still above the baseline. The threshold also falls because, at higher prices, low-income-high-ability students are less likely to enroll. Fewer students are crowded out on the admissions margin from the policy change, but more choose not to attend because it is not worth the cost. From the enrollment and price changes, the college raises more revenue for instruction, improving quality and thereby increasing demand and the entire distribution of prices further. Moreover, with less crowd out of high-income-low-ability students, the spillovers on the other privates are smaller.

The change in attendance patterns and quality from increases in private college sticker prices has an ambiguous effect on the labor market. More students are admitted to the highest quality college and the privates collect more tuition revenue to improve quality but fewer choose to enroll at higher prices. We find that increases in private sticker prices decrease total human capital in the economy. Therefore, the college wage premium rises with private market power.

3.5.3 Equilibrium Impacts of Sticker Price Increases

As highlighted in Section 3.4, the sticker price elasticity's effect on private college admission has important implications for who benefits and who bears the burden of public investment in higher education. The more the sticker prices increase, the more the privates can increase their market power and modest-ability students can compensate for their negative peer effects. The elasticity therefore determines who has access to the highest quality private after a 10 percent decrease in appropriations per student ($z = 0.9\bar{z}$).

Panel (a) of Figure 3.2 plots the change in HP college's minimum ability threshold as a function of the sticker price elasticity. As it increases in magnitude, the minimum ability threshold falls. However, the elasticity would need to be less than -0.325 for the admission threshold to be lower than the baseline, which is well outside the confidence interval of our estimate from Table 3.2, shown by the virtual dotted lines. After the decrease in public investment, the highest quality private becomes more selective and some modest-ability students are crowded out, although the number depends on the sticker price elasticity.

Increases in private college sticker prices also affect attendance by pricing lower income households out of the market. Those students either choose to enroll in the lower quality public colleges or do not attend any. Although the privates use the extra revenue to improve their own quality and increase their students' human capital, Panel (b) of Figure 3.2 implies that aggregate college human capital decreases after the cut to state spending and is decreasing with the magnitude of elasticity. Increases in private college prices, and therefore market power, exacerbate the college wage premium. Increases in the college wage premium also drive up demand for the privates, allowing them to charge higher tuition, although they are limited by their sticker prices. These two effects build on one another, but the extent to which they do depends on the elasticity.

3.6 The Incidence of Public Investment

We consider several different methods of measuring who benefits and who bears the burden of public investment in higher education. We begin by showing that the model replicates the main findings of the accounting approach. We then present the model implied CV measure in two steps to illustrate the general equilibrium mechanisms. First, the private college sticker prices are held fixed. In this exercise, private college market power increases but only to a limited extent. Then, sticker prices adjust based on the results from the previous section to better understand the role of private college market power. In both CV exercises, the labor market returns to college endogenously adjusts as a function of public investment and private market power. Appendix C.3 additionally presents comparative statics to further illustrate the key mechanisms of how changes in the college market affect incidence.

3.6.1 The Model Implied Accounting Approach

The model replicates Pechman's and Bill Johnson's accounting approach findings that spending is weakly progressive. Panel (a) of Figure 3.3 presents the distribution of net benefits by parental income, that is, instructional expenditure plus capital costs minus tuition and taxes paid. Each open circle represents the average net benefits for a household of type (b, y).[22]

Based on the model implied accounting approach, the very low income do not benefit since they are unlikely to attend given the costs even with the maximum amount of federal financial aid. Slightly higher income households still receive the maximum federal aid at both public colleges but tuition is less of a financial burden, leading to large attendance rates and a positive mass around $40,000. As income increases, aid for the in-state public falls while households still receive the maximum amount at the out-of-state option and the privates. Net benefits decrease sharply since they are less likely to attend their in-state public college. High income households have large negative net benefits since they are much more likely to attend the private colleges but pay high income taxes.

Panel (b) of Figure 3.3 shows a weak, but negative, association between ability and net benefits, or that lower ability students benefit more from public investment. The in-state public is the best option for these students since they are not admitted to the highest-quality private college, and have relatively high effective marginal costs and prices at the others. Higher ability students receive merit aid at the privates, which makes those higher quality colleges more affordable. Since income and ability are positively correlated, average taxes paid are also increasing in ability. Taken together, there is a negative association between net benefits and ability for higher ability students.

We additionally calculate the accounting measure for a 10 percent decrease in state appropriations as the change in instructional expenditure minus the change in tuition and taxes paid to support the public colleges, holding college attendance decisions fixed:[23]

$$(I(z) - I(0.9z)) - (t(z) - t(0.9z))y - (T(z) - T(0.9z)) \tag{3.18}$$

[22] The figures show the benefits by income and ability, averaging over 20 sets of idiosyncratic preference draws, to emphasize the distributional effects of public investment. Appendix C.2 additionally presents the accounting approach for each household-enrollment decision (b, y, ϵ).

[23] Note that since attendance decisions are fixed, so are capital costs.

The decrease in appropriations has a direct effect on the college and state budget constraints, affecting both instructional expenditure and tax rates. Public colleges also increase tuition to recover some of the lost revenue. Some lower income households also receive additional federal financial to offset the tuition increase. This equation is analogous to Equation 3.1, which allows us to compare the two approaches for measure incidence and understand the bias from ignoring the general equilibrium effects of public investment.

The distributional effects for a 10 percent decrease shown in Panels (c) and (d) are qualitatively similar to the total net benefits. However, the share of the net benefits that go to low income household shrinks because they now pay for a larger share of their lower quality education. Additionally, slightly higher income households see no change in relative net benefits because new federal financial aid largely offsets the increase in tuition. Higher income households also have a similar share of net benefits.

3.6.2 The General Equilibrium Incidence of Public Investment

We now reexamine incidence using our model and the compensating variation framework described in Section 3.2, since the accounting approach misses important general equilibrium consequences of public investment in higher education. We first decrease state appropriations per student by 10 percent, or about $700, holding private college sticker prices fixed. This counterfactual corresponds to the equilibrium presented in Column (2) of Table 3.3. Panels (a) and (b) of Figure 3.4 present the distribution of CV by income and ability, respectively. A positive number indicates that a household is better off when state spending is high; she would need to be given a positive amount of money to be just as well off after the decrease in appropriations per student. Each open circle represents the average CV for a household of type (b, y) based on 20 draws from the preference distribution. We focus on the average CV to highlight the distributional effects.[24]

Public investment in higher education is weakly progressive, but high-income-modest-

[24] Appendix C.2 presents the results conditional on a household's attendance decision based on the preference draws.

ability students benefit the most. They are crowded out of the highest quality private since more low-income-high-ability students now enroll. The implied change in the minimum ability threshold generates the large positive spike in CVs in panel (b) near the 60th percentile of the ability distribution. Many of these students lose access to their preferred college and face a large welfare cost to move to their next best option. The negative welfare effects from changes in admission are especially concentrated among the high-income because low-income-modest-ability households are less likely to enroll in the high price privates where they are charged the fill sticker price, and even those that do are not as harmed when moving to the lower price public colleges. That is, given relative prices and qualities, the publics and privates are closer substitutes for lower income households.

Most students who enroll in public colleges when state support is high are also negatively affected by the appropriations decrease, but the welfare effects are smaller in magnitude than those on the high-income-modest-ability. This mainly affects above median income households who are the most likely to enroll in the in-state public, as seen in the small hump in panel (a) from \$50K to \$100K. Their preferred baseline ($z = \bar{z}$) college option becomes lower quality and more expensive, implying a positive CV, especially for those that still enroll. They are not very mobile across sectors because the privates are too expensive and the returns to any college are high, and so they bear some of the burden of the policy change. However, based on our parameterization and calibration, some lower ability students who still enroll in the public colleges after the appropriations decrease are better off because the value of the increased returns to college outweighs the decrease in public quality. By missing the general equilibrium labor market effects, the accounting approach overstates the negative effects on these students. Slightly lower income households are less affected because their federal financial aid increases with prices.

Public college students can partially avoid the negative effects of higher tuition and lower college quality by enrolling in the privates or not enrolling in any college. However, when they leave the public sector after state divestment, they still face the general equilibrium effects on the private and labor markets. Changes in college choice, as a function of state funding, are important for assessing who benefits and who bears the burden. The accounting approach misses both of these margins. The low income are especially harmed because the price increases push some of them out of the college market, and they then see their wages fall from the decreases in both human capital and the returns to human capital as non-college workers. The higher income, who are less burdened by the private price increases, are more likely to shift to the higher quality colleges where they benefit from the increased returns to college. The overall welfare effect is ambiguous in sign for these students, but since they no longer pay public college tuition to help finance public expenditure, CV is often negative.

At the other extreme, high-income-high-ability students benefit the most from state divestment in public higher education. As more high-ability students shift to the HP college and all privates charge higher average prices from the increased aggregate demand, the colleges become higher quality. However, since these households are high-income, they always pay the sticker price. They receive a higher quality education, when the returns to quality are higher, while paying the same tuition and lower taxes. Although missed by the accounting approach, the general equilibrium consequences of public investment have large and varied effects on the high-income. Overall, excluding these effects biases incidence in both directions.

The equilibrium changes in the private and labor market from a decrease in state appropriations have an ambiguous welfare effect on lower income baseline private college students. The moderate income households value the increased quality and returns to college more than the added tuition costs, while the reverse is true for lowest income. The higher prices even lead some students to enroll in the less expensive and lower quality public colleges or not enroll in any college– increased private market power crowds out some low-income students.[25] For those who switch to the public sector from the privates, the welfare effect could be positive or negative, depending on the individual's benefits from the increased returns

[25] Very few students are no longer able to afford college from the price change. However, price increases enough that the privates are no longer their best option.

to college. However, CV is shifted up because they only pay public college tuition when spending is low.[26]

Finally, the equilibrium labor market changes have a negative welfare effect on baseline non-college workers. Their wages decrease as college human capital becomes more scarce from the increased private college market power and decrease in public college value. Therefore, the changes in the college market spill over to non-college workers since the two are linked through the economy's production function– with more public investment, non-college workers benefit from less competition in the labor market. While the accounting approach implies that these students have negative net benefits from paying taxes, our approach additionally captures the added value from labor market spillovers. Ignoring these general equilibrium effects leads the accounting approach to understate progressivity since these households are more likely to be low-income.

Inframarginal college attendees in the baseline enroll in college after the shock because of the decrease in relative earnings. The welfare effect is ambiguous in sign, similar to what we describe for several other cases above, because both price and the returns to college increase over the baseline. This trade-off is more likely to be a net benefit for the high-income baseline non-college workers, who shift to colleges 3 and 4. Lower income, modest ability students switch to the publics to avoid the low non-college wages and are worse off but often have positive CVs because the policy change greatly increases how much they pay to help finance public expenditure with tuition.

With increased demand for private colleges after the negative shock to state appropriations, the privates will increase their sticker prices. Panels (c) and (d) of Figure 3.4 present the distribution of CV allowing the maximum prices to adjust based on our estimates from Table 3.2, as in Column (3) of Table 3.3. Since the sticker prices minimally increase, the results are qualitatively similar to the previous case. However, the mass of positive points for households in the top of the income distribution in Panel (c) is slightly smaller than in

[26] In practice, the difference between ν and CV in Equation 3.1 is small because the low-income receive federal aid to cover their tuition expenses.

Panel (a). For example, in panel (c), 1.03 percent of households in the top 10 percent of the income distribution have positive average CV. In panel (a), the corresponding percent in 1.21. With higher sticker prices, more students are able to sufficiently compensate the high quality private for their modest ability. Public investment in higher education is still weakly progressive, but the largest benefits still go to modest ability students who only have access to high quality private when state spending is high.

Sticker price increases also create other equilibrium changes in college price and quality, the wage premium, and taxes that matter for incidence. Overall, these effects on individuals are much smaller in magnitude than changes in admission, but impact all households. The sign of these effects on welfare varies importantly with income and ability, but the implications for incidence are analogous to the fixed sticker price case because they all move in the same direction from the baseline. Compared with the fixed sticker price case, high-income-high-ability students receive much of the new benefits because HP college's quality increases along with the returns to quality. Since they are high income, they always pay (nearly) the full sticker price so tuition only increases modestly. Most of the new burden from the sticker price increases falls on the lower income, who are more harmed by the increase in prices than the increase in private college quality or are even priced out of the private market.

Taking the results from the two counterfactuals together, we find that high-income-modest-ability students benefit the most from public investment in higher education because the general equilibrium effects on the private college and labor markets are large. The externalities on these students from changes in the private market are smaller when the privates raise their maximum tuition, but colleges only minimally increase their prices after state divestment. Since these households are very likely to attend the privates under the current level of state appropriations, $\bar{z}$, the accounting approach would instead say that they do not benefit but pay high taxes. However, these households only have access to the highest quality private college when state spending is sufficiently high. By ignoring the impacts of public investments on other parts of the economy, the accounting approach mischaracterizes incidence.

3.6.3 The Bias from Ignoring Equilibrium Effects

Since the accounting approach ignores the general equilibrium effects of public investment on college enrollment, and the private college and labor markets, it mischaracterizes who benefits and who bears the burden. To further highlight the distributional implications, we consider two policy relevant measures of the bias. First, in Panel (a) of Figure 3.5, for each income-ability decile, we present the average CV measure in dollars with endogenous sticker prices and labor market returns to college minus the accounting approach for a 10 percent decrease in state appropriations (Equation 3.1 - Equation 3.18). In Panel (b) we define the bias as the percent of households with a positive CV minus the percent with a positive accounting measure. More negative values are shaded lighter and indicate that the accounting approach overstates the benefits by more.

Overall, the accounting approach overstates both the average benefits and the number of households that benefit from public investment in higher education, but the magnitude varies importantly across the income and ability distribution because the general equilibrium effects are large and ambiguous in sign. The largest differences in magnitude in Panel (a) are for the highest income decile households. They benefit from the general equilibrium implications of state divestment on private college quality and the labor market returns to college. Ignoring these effects leads to substantial bias for measuring the average benefits. The bias is large enough that the accounting approach also gets the sign of the welfare effect wrong for about 16 percent of the highest income decile households.

The average net benefits and CV for the very high income households in the seventh ability decile (between the 60th and 70th percentiles) are much closer in magnitude but only because the welfare effects from changes in private college market power vary in sign. The accounting approach misses several important equilibrium effects that act in opposite directions. High-income-modest-ability students between the 60th and 61st ability percentiles are crowded out of the highest quality private when state funding decreases and are much worse off, while those above the 61st benefit from the increases in private quality and the returns to college.

Slightly lower income households in the seventh, eighth, and ninth income deciles, are similarly affected, although to a lesser extent, from the increase in private market power and the returns to college. Increased private college market power leads to price increases for these students since they are less likely to pay the full sticker price when state funding is high. These households additionally have more negative CV, on average, because they are especially likely to switch to the private colleges after the public funding decrease. This allows them to avoid the negative effects of divestment on the public colleges, and they no longer pay public college tuition. Those who remain at the public colleges also benefit from the increased returns to college human capital. These three equilibrium effects are missed by the accounting approach.

The difference between CV and the accounting measure is nearly uniformly decreasing in ability for the relatively high income, except for the high-income-modest ability who are crowded out of the highest quality private college. Both the human capital returns to college quality and the labor market returns to college are higher for higher ability students. However, the same pattern does not hold for the percent with positive benefits because public college enrollment and the likelihood of moving from a public to private college after state divestment are non-linear functions of ability, conditional on income.

The interaction between college choice, private college market power and the general equilibrium labor market affects of state divestment are especially important for households in the fifth and sixth income deciles. Many of these households only enroll in public colleges when state spending is low. The private colleges become too expensive, and non-college wages fall. The large changes in tuition from changes in college choice not captured by the accounting approach have a large impact on CV. The accounting approach therefore understates the percent of high-ability households in the sixth income decile who benefit from public investment.

These interaction effects are also important for slightly lower income households in the third and fourth deciles. Since the cost of college imposes a large financial burden, the public college price increases lead some of them to no longer enroll. They are made worse-off from the decrease in relatively wages and loss of human capital. However, they only pay public college tuition when state spending is high, which pushes CV down.

Finally, although the difference between the two approaches to measure incidence is small for households in the bottom two income deciles, the composition of the total effects are different. These households are mainly affected by the tax changes, which both approaches capture, as well as the general equilibrium labor market effects. Since fewer students enroll in the public colleges after the spending decrease, states and the federal governments collect less tax revenue. The accounting approach understates the corresponding decrease in equilibrium tax rates because student enrollment decisions are fixed. This alone would lead CV to be larger than the accounting measure. However, for these students, the accounting approach overstates the total benefits because it does not include the negative labor market externalities on mainly low-income non-college workers.

3.7 Conclusion

Previous measures of the incidence of public investment in higher education implicitly ignore general equilibrium effects. We use a compensating variation framework in a general equilibrium model of higher education to understand how responses in the private college and labor markets affect the distribution of benefits. We show that public investment puts competitive pressure on the privates and reduces the college wage premium, which creates benefits for students who do not currently attend public colleges. Unlike the previous accounting approaches, our measures capture the effect of public investment on private college prices and admissions decisions, student enrollment decisions, and the aggregate implications for relative wages.

Building on a class of stylized models of higher education, we find that these general equilibrium effects are large in magnitude and vary importantly with both income and ability.

This makes signing the bias for overall progressivity from the accounting approach difficult. Students who look similar based on their income and college attendance decisions at the current level of state spending can face different consequences of state divestment that matter for incidence. For example, in our model, high-income-modest-ability students benefit the most from public investment because they are only admitted to the privates when the publics provide sufficient competitive pressure, that is, when state support is high. However, the high-income, slightly higher ability students are not crowded out after a negative appropriations shock. They retain access to the highest quality private colleges when increased private market power improves college quality and raises their wages.

While the role of private colleges and the labor market enhance our understanding of incidence in this setting, both approaches focus on the private returns and cost. Higher education impacts the distribution of human capital and therefore the tax base as well as the demand for other public expenditure, like assistance programs. If the social benefits mainly impact lower income households, then both measures understate the progressivity. If the increase in human capital leads to lower tax rates on the wealthy, then we would understate the regressivity. Further work is needed to understand how the social returns to public investment affect incidence.

Table 3.1: Calibration Parameters

Parameter	Name	Value
Households:		
α_p	Parent Utility Shifter	20.832
α_s	Student Utility Shifter	25.320
$w\beta_1^c$	College Wage Shifter	2.614
$w\beta_2^c$	College Wage Ability Share	0.505
$w\beta_3^c$	College Wage Subst. Parameter	-0.936
$w\beta_1^h$	Non-college Wage Shifter	2.359
$w\beta_2^h$	Non-college Return to b	0.111
$w\beta_3^h$	Non-college Return to y	0.138
L	Living expenses	\$5479
b	Ability Distribution	$ln(b) \sim N(1.0, 0.0225)$
y	Income Distribution	$ln(y + 42470) \sim N(11.500, 0.298)$
$corr(b, y)$		0.359
Governments:		
$\overline{T}_s$	Resident; Non-Resident Tuition	\$7113; 19345
$\overline{z}_s$	Per Student Subsidy	\$7062
EFC	Effective Family Contribution	$\max\{0, .48y - 10300, .69y - 22500\}$
$\bar{A}$	Maximum federal aid	\$4000
$\kappa_s; \kappa_{so}$	In/Out Tuition Elasticity	-0.263; -.066
Colleges:		
ϕ	Private Productivities	1.089;1.118;1.182
γ	Return on Peer Quality	0.049
ω	Return on Instruction	0.065
E_j/k_j	Baseline Endowments/Student	\$221, 603, 3850
P_j^c	Baseline Sticker Prices	\$24082; 31724; 34118
κ_p	Sticker Price Elasticity	-0.126
Costs:		
F	Fixed Cost	\$231
$v_1; v_2$	Custodial cost parameters	\$0; 18057
Labor Market:		
ρ	CES Substitution Parameter	0.412
s	CES College Share	0.461

Table 3.2: The Effect of State Appropriations on Private College Sticker Prices

	(1)	(2)	(3)
	Log Price	Log Price	Log Price
Lower quartile	-0.003	-0.015	-0.042
	(0.035)	(0.035)	(0.054)
Lower-middle quartile	0.020	-0.002	-0.031
	(0.025)	(0.026)	(0.049)
Upper-middle quartile	-0.069*	-0.093***	-0.126**
	(0.038)	(0.036)	(0.057)
Upper quartile	0.033	0.011	-0.024
	(0.028)	(0.025)	(0.052)
Observations	14705	14705	14705
Year FE	Yes	Yes	No
State Linear Time Trend	No	Yes	No
State-by-Year FE	No	No	Yes
Institution FE	Yes	Yes	Yes

Note: Each column presents the coefficients from a regression of the log of institution relevant appropriations interacted with quartiles of SAT scores on the log of private college sticker prices using a balanced panel of institutions from 2000 to 2017 with various institution and time controls. Standard errors clustered by institution. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.3: State Investment and the Higher Education Market

	(1) $\overline{z}_s$ with P_j^c	(2) $0.9\overline{z}_s$ with P_j^c	(3) $0.9\overline{z}_s$ with $(1-0.1\kappa_p)P_j^c$
Enrollment:			
Publics	0.131	0.129	0.130
3	0.041	0.041	0.040
4	0.032	0.032	0.032
5	0.031	0.032	0.031
Expenditure per Student I_j (\$10K):			
Publics	1.044	0.993	0.993
3	1.551	1.619	1.630
4	2.023	2.086	2.098
5	2.988	2.995	3.032
Mean Student Ability θ_j (percentile):			
Publics	65.017	65.099	65.147
3	79.838	79.560	79.765
4	83.184	83.012	83.182
5	89.965	90.139	90.007
Quality q_j:			
Publics	1.056	1.052	1.052
3	1.181	1.184	1.185
4	1.237	1.240	1.240
5	1.346	1.346	1.347
Minimum Ability Enrolled (percentile):			
Public Resident	0	0	0
Public Non-resident	0	0	0
3	0.008	0.008	0.008
4	0.008	0.008	0.016
5	60.061	62.331	61.443
Average Tuition (\$10K):			
Publics	0.751	0.769	0.769
3	2.153	2.233	2.251
4	2.728	2.799	2.817
5	3.393	3.400	3.443
Percent Increase in Baseline Wages Per Unit Human Capital			
w_c	-	0.239	0.253
w_h	-	-0.178	-0.188
Tax rates			
t_s	0.025	0.022	0.022
t_f	0.005	0.005	0.005

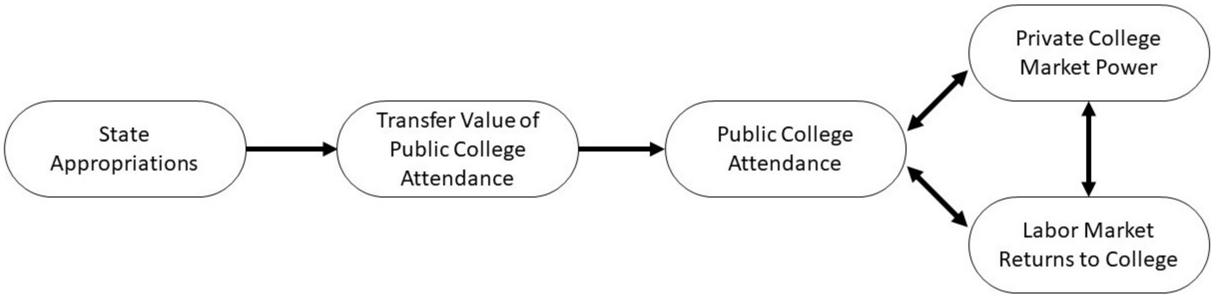
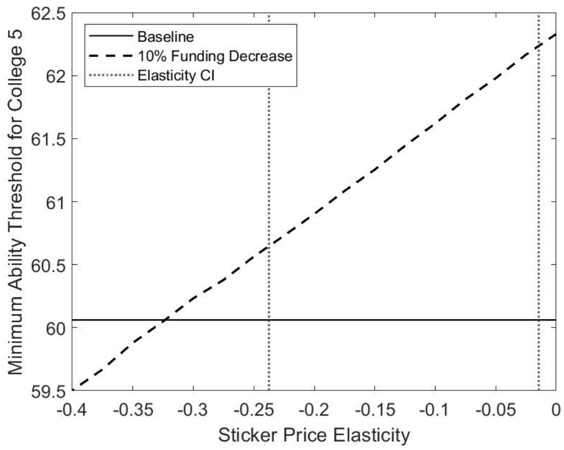
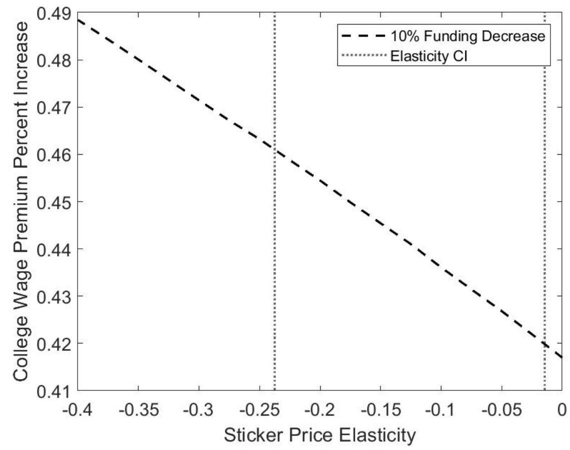


Figure 3.1: General Equilibrium Implications of State Appropriations



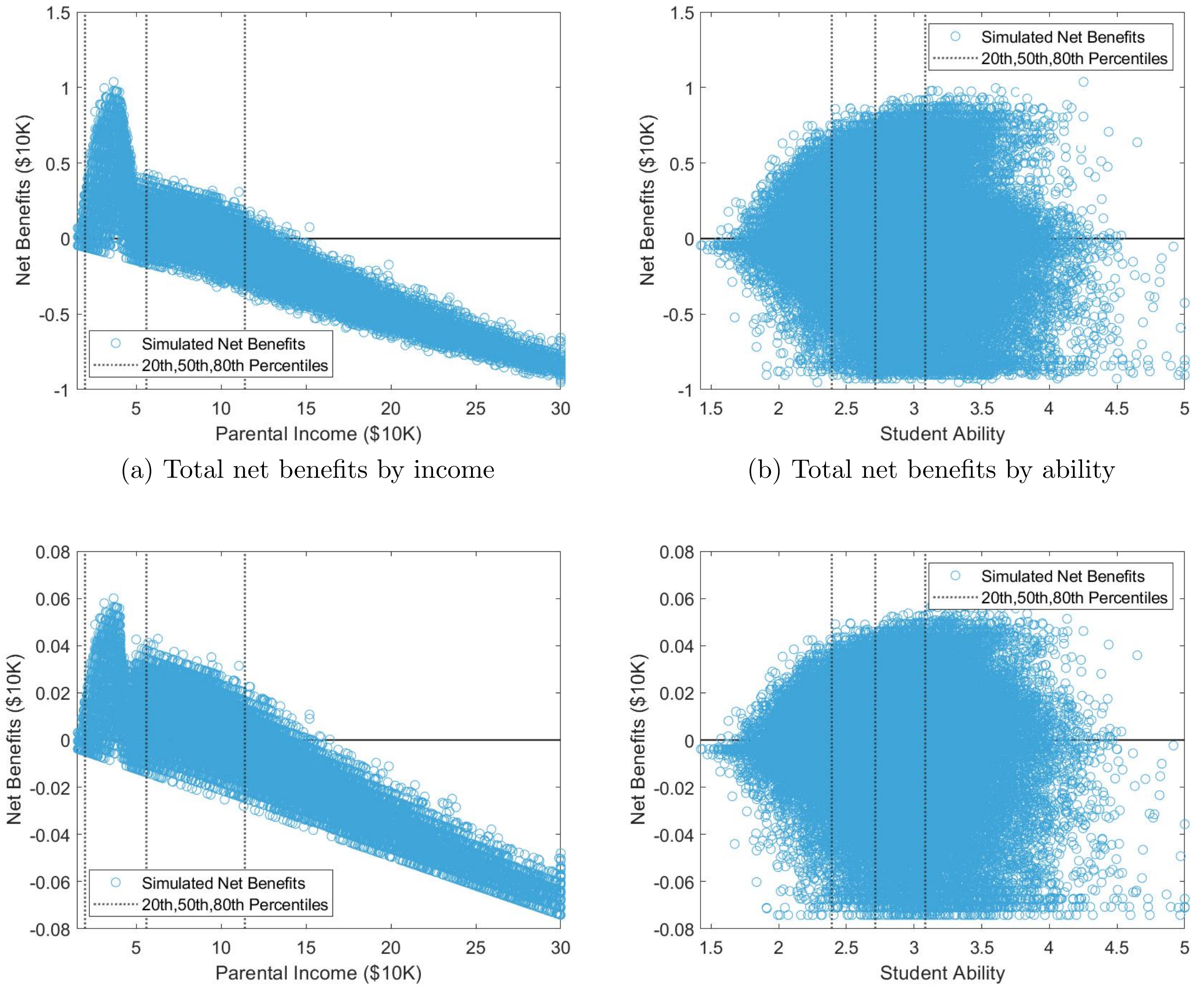
(a) College 5 Admission



(b) College Wage Premium Increase

Figure 3.2: Equilibrium Impacts of Sticker Price Increases

Note: Panel (a) plots the relationship between the sticker price elasticity and college 5's minimum ability threshold for a 10 percent decrease in appropriations per student. The horizontal solid line shows the baseline admission's threshold for reference. Panel (b) plots the college wage premium, w_c/w_h, for the same decrease in funding as a function of the elasticity. In both panels, the vertical dotted lines show the upper and lower bounds of the elasticity's confidence interval from Table 3.2.



(a) Total net benefits by income

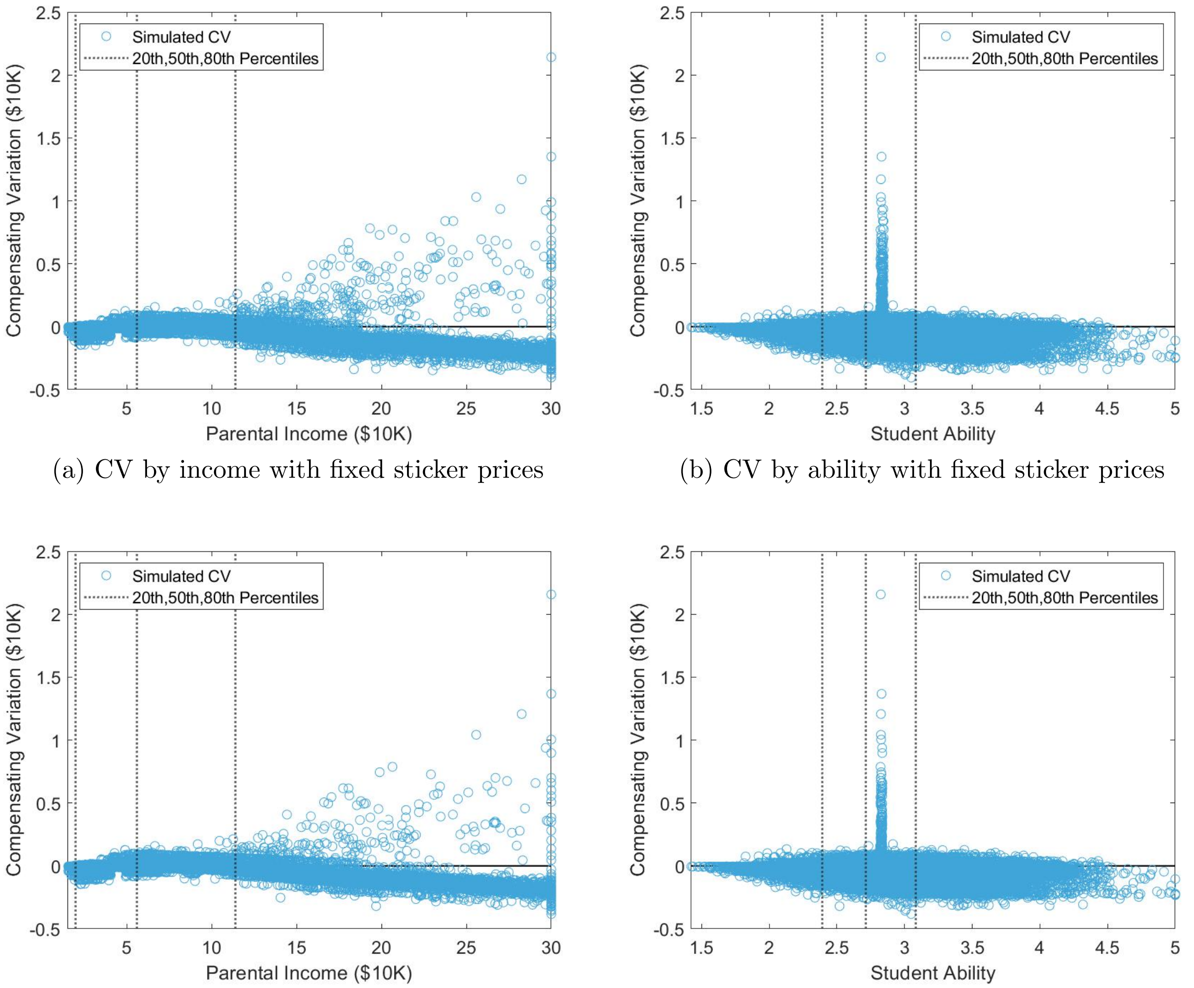
(b) Total net benefits by ability

(c) 10% decrease net benefits by income

(d) 10% decrease net benefits by ability

Figure 3.3: Net Benefits for Public Higher Education Investment

Note: Each open circle represents average net benefits for one draw from the joint distribution of income and ability using the accounting approach. For each household, we take 20 draws from the preference distribution, find the implied college enrollment decision, and then average over the corresponding net benefits. Panels (a) and (b) present the distribution for public colleges, while (c) and (d) for a 10 percent decrease in appropriations. The dotted vertical lines represent the 20th, 50th and 80th percentiles of the income and ability distribution. A positive value implies that the average household of type (b, y) receives more in college expenditure than it pays in tuition and taxes.



(a) CV by income with fixed sticker prices

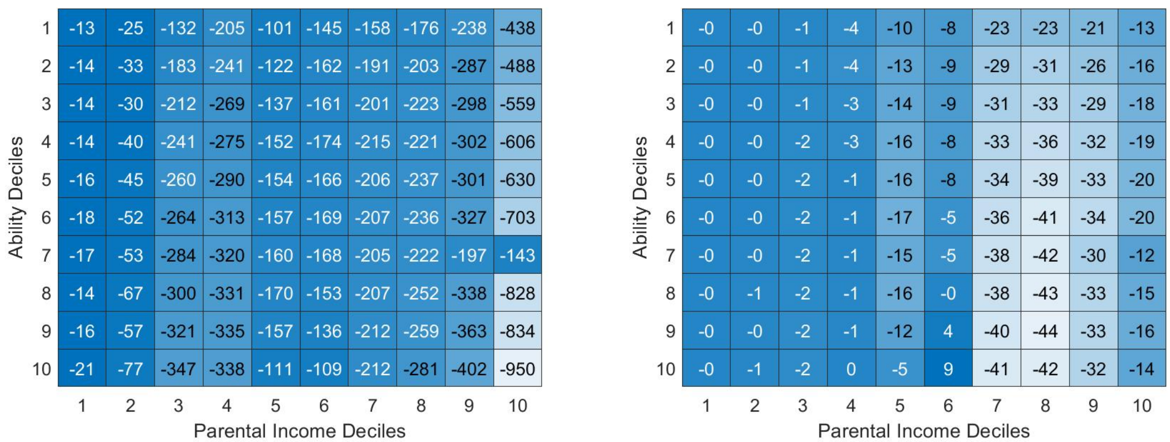
(b) CV by ability with fixed sticker prices

(c) CV by income with endogenous sticker prices

(d) CV by ability with endogenous sticker prices

Figure 3.4: Compensating Variation for Public Higher Education Investment

Note: Each open circle represents average compensating variation (CV) for one draw from the joint distribution of income and ability. For each household, we take 20 sets of draws from the preference distribution, compute the corresponding CV based on equilibrium changes and the enrollment decisions, and then average over the 20 estimates. The dotted vertical lines represent the 20th, 50th and 80th percentiles of the income and ability distribution. A positive value in panels (a) and (b) denotes that the household is better off under the baseline than when appropriations decrease by 10 percent, holding private college sticker prices fixed. Panels (c) and (d) allow private colleges' sticker prices to endogenously adjust.



(a) Average Benefits

(b) Percent Positive Benefits

Figure 3.5: Difference Between CV and Accounting Approaches

Note: Panel (a) shows the difference between CV and 10 percent of net benefits from the accounting approach in dollars by income-ability deciles. Panel (b) shows the difference in the percent of households with positive compensating variation minus percent with positive net benefits by income-ability deciles.

Appendices

Appendix A

Appendix to Chapter 1

A.1 Service Plan Example

Before the failed Pflugerville annexation in 2017 the service plan included (Austin Community College 2020):

- A workforce innovation campus/center to meet workforce needs in healthcare, IT, and/or advanced manufacturing and transfer education courses in PfISD
- Student support services including advising, admissions and records, academic counseling, career counseling, financial aid, and student life
- Customized training to assist businesses and organizations in their employee training; access to degrees and/or certificates in transfer, career pathways, and Continuing Education non-credit
- Access programs such as English-as-a-Second Language (ESL), GED, and Adult Basic Education to prepare people for entering college credit programs
- Credit or non-credit entry-level, and re-entry job-training programs that can be completed in one year or less
- Eligibility to vote in elections of the ACC Board of Trustees

A.2 Data Appendix

This appendix describes each data source used in the structural model.

A.2.1 Housing Prices

Data on housing prices consists of two sources: (1) Repeat-sale home price indices estimates from Bogin et al. (2019); (2) CoreLogic Tax Roll Record File. The first source is used in the event-study analysis and together with the second to estimate the structural model.

1. Home price indices for the event-study analysis and public good production function estimation: Bogin et al. (2019) estimate repeat-sale home price indices for Census Tracts in the United States. These estimates net out the characteristics of the home, like the number of bedrooms, and provide a measure of the relative price of land in a Census Tract over time. Census Tracts are mapped to school districts using information provided by the Missouri Census Data Center.

2. Community prices in the structural model: The cost of living in community j and receiving public good g_j is determined by the price of a unit of housing services, p_j. p_j is estimated in 2015 using a hedonic regression framework following Sieg et al. (2002, 2004). This process nets out the structural characteristics of the home to capture the value of living in a community, rather than the correlation between household income and the size of the home.

   Sieg et al. (2002, 2004) show that community prices in locational equilibrium models can be recovered by regressing the rental value of housing on home characteristics and community fixed effects. Their estimation framework is implemented in this setting using the CoreLogic Tax Roll Record File from 2016. An observation in the CoreLogic data is a parcel of land, for example a single family home, and contains information on the characteristics, market value, and assessed value. The sample includes single family homes sold during or after 2013. Since the data do not have income information for the owners, it is not possible to directly compute the rental value of housing, V_r, from the market value and instead the estimation strategy includes year fixed effects to capture i and ζ and control for the property tax rate t_j. Table A.1 presents the estimates. The model also requires that prices are strictly increasing in community income rank. Therefore, after estimating prices in each community-year, they are smoothed using least squares splines, enforcing that requirement.

A.2.2 Amenities

There are several data sources used to compile the set of amenities for each Census Tract.

1. College spending: Community college instructional expenditure per student comes from the Integrated Postsecondary Education Data System (IPEDS). It is available beginning in 1987. IPEDS also provides the coordinates of the college campus.

2. Distances: Distances to the community colleges' main campuses are computed using the geodist command in STATA based on latitude and longitude. Census Tract centroid latitudes and longitudes are provided by the Missouri Census Data Center.

3. ISD spending: K-12 expednitures per student data comes from the Elementary/Secondary Information System (ElSi), part of the National Center for Education Statistics (NCES). The data is first available for the 1989-1990 academic year. It is next available for 1991-1992. Beginning in 1994, the data is reported every year until 2016-2017. Census Tracts are matched to ISDs using the mapping files from the Missouri Census Data Center.

4. Crime rate: Violent and property crime rates by police precinct for 1986-2014 come from the Uniform Crime Reporting Statistics (UCR). After 2014, data come from the Texas Department of Public Safety. Census Tracts are matched to police precincts to minimize the distance from the center of the Census Tract using the geonear command in STATA. Police precinct latitudes and longitudes are provided by the Law Enforcement Agency Identifiers Crosswalk.

A.2.3 Household Information

The model uses several pieces of information on households: (1) residency choices; (2) income; and (3) home values. Data is collected for the 39 school districts that are at least partially in the Austin metropolitan statistical area, consisting of Bastrop, Caldwell, Hays, Travis, and Williamson counties.

1. Residency choices: The American Community Survey (ACS) data are available at the school district level from IPUMS. The number of households in each school district gives the population or residency choices. The share of each community is then computed as the percent of the households in each school district.

2. Income: The ACS reports the 20th, 50th, and 80th percentile of the household income distribution for each ISD.

3. Rental Values: The ACS contains data on the binned distribution of home values for owners and rental values for renters. Poterba (1992) describes how to convert home values V_h to yearly rents V_r: $V_r = [(1 - t_y)(i + t_j) + \zeta]E[V_h|y = y_i]$, where t_y is the marginal tax rate for household with income y, i is the nominal interest rate, t_j is jurisdiction j's property tax rate, and ζ is the sum of the risk premium for housing investments and the annual percentage rates of depreciation and maintenance less the annual rate of appreciation of house value.

   From Poterba (1992), the risk premium is 0.04, maintenance is 0.02, and depreciation is 0.02. Following Epple and Sieg (1999) and the subsequent literature, t_y is 0.15 since the data only provide the marginal distributions of income and housing values. While income taxes are progressive, higher income households are more likely to earn income from sources with lower tax rates, like capital gains. The 5 year average of CPI inflation for Houston in 2015, the closest metro area to Austin for which the BLS data are available, is 0.018. The nominal interest rate on fixed rate mortgages is 0.039 in 2015 based on estimates from Freddie Mac. Finally, I compute $E[V_h|y = y_i]$ from the binned ACS data.

   The tax rate t_j for each ISD is based on the sum the school district tax rate and the community college tax rate for those in the taxing district since those are the two most important rates for households decisions. Households also face taxes from the county and other special districts but these vary within school districts. To account for these, the sum of ISD and CC rates are scaled by $\frac{1}{.55}$ since the Texas Comptroller notes that those two taxes comprise about 55 percent of the total property tax rate, on average. After converting home values to rental values, I combine them with the data on renters and compute the 20th, 50th, and 80th percentiles of the implied rental distribution.

A.3 Additional Impacts of Annexations

The effects on aggregate home values and the home price indices in Figure 1.1 are intended to summarize the benefits to households from joining the tax base. However, even though aggregate home values increased, policy makers may be concerned that the increased tax rates to fund the public good induces some households to leave. To proxy for the effect on migration, Figure A.1 estimates the effect of joining the tax base on the number of single family homes. Unfortunately, this data is not available for 1994, which limits the ability to properly test for pretends since 28 of the annexations occur between 1995 and 1998. However, the pattern suggests that joining the tax base did not decrease the total number of households in the community.

Figure A.2 plots the effect of joining the taxing district on K-12 education spending per student to understand how it impacts other locally provided public goods. The estimates show no effect of annexation on spending.

A.4 Impact on Firms

While the focus of this paper is on household decision-making in the presence of spillovers, firms are an important part of the property tax base to fund the public good. In 2016, firms comprised 12 percent of total property value in the recently annexed communities, on average. Firms in the taxing district may have access to higher income customers and a more educated workforce, but face higher taxes. Since customers and workers are mobile within the surrounding area, the benefits from being in the tax base are likely quite small for firms. If firms exit or no longer enter after annexation, total property values can decrease, making it more difficult to finance local services. To understand the effect of joining the tax base on business activity, I estimate Equation 1.1 with the log of commercial firm property value as the dependent variables and plot the coefficients in Figure A.3. The point estimates suggest that there is no relationship between annexation and firm property value.

Table A.2 presents the impact on other communities to understand the spatial effects. Column (1) shows a small positive but imprecise effect of the annexation on firms in the original tax base. Column (2) shows the estimated effect on the remaining part of the service area who still do not pay property taxes to support the community college. These communities benefit from having relatively lower tax rates and therefore the annexation leads them to have increased firm property value. Overall, these results suggest that increased centralization leads to more firm activity that is concentrated outside of the annexed community, which provides additional evidence that the tax base is too small. Voters in a community considering annexation may not internalize these benefits to others.

A.5 Measuring Spillovers Decomposition

This section presents summarizes the partial and general equilibrium per capita welfare estimates from a 10 percent increase in Austin Community College's expenditure per student, financed by an increase in property taxes to those in the taxing district.

Table A.1: Price of Housing Services Hedonic Regression

	Log Sale Price
Log Lot Size	-0.086
	(0.177)
Log Lot Size Squared	0.010^{+}
	(0.005)
Log Building Size	0.300
	(0.207)
Log Building Size Squared	0.039^{*}
	(0.016)
Log Age	0.019
	(0.056)
Log Age Squared	0.048^{+}
	(0.026)
Log Age × Log Lot Size	-0.008
	(0.008)
Log Age × Log Building Size	-0.021^{*}
	(0.009)
Log Lot Size × Log Building Size	-0.008
	(0.015)
Log Bedrooms	0.011
	(0.037)
Log Bathrooms	0.050^{*}
	(0.022)
$\mathbb{1}$ Has Fireplace	-0.023
	(0.040)
$\mathbb{1}$ Has Pool	0.148^{***}
	(0.016)
Log Property Tax Rate	-0.829^{***}
	(0.020)
Observations	110670

Note: The table presents the estimates from a hedonic regression of the log of home sales on home characteristics. For sales with missing values for any variable, the value is set to 0 and a indicator for having a missing value is included. The regression additionally includes year and ISD fixed effects. Standard errors are clustered at the ISD. The sample is limited to homes sold since 2013. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.2: The Effect of Annexation on Firm Activity in Other Communities

	Initial Tax Base	Remaining Service Area
	(1)	(2)
	Property Value	Property Value
Log # Other School	0.265	0.231^{+}
Districts in Taxing District	(0.281)	(0.129)
Observations	442	2158

Note: Column (1) shows the estimated effect of annexations on communities that were part of the tax base in 1991. Column (2) estimates the effect on the remaining communities in the service area. All specifications include year and ISD fixed effects. Standard errors are cluster bootstrapped with 999 replications, clustered by school district. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

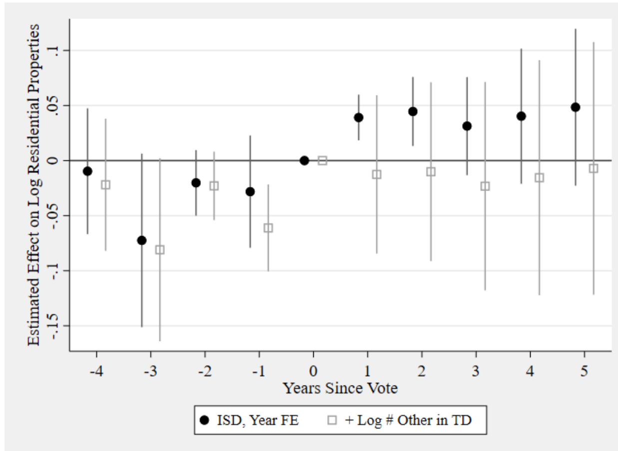


Figure A.1: The Effect of Annexation on the Number of Single Family Homes

Note: The figure plots the event-study coefficients and 95 percent confidence intervals for the effect of joining the tax base on the log of the number of single family homes in the school district using the sample of school districts that are annexed between 1991 and 2016. Standard errors are cluster bootstrapped with 999 replications, clustered at the school district level. The solid black circles are from Equation 1.1, and the clear squares additionally control for the log of the number of other school districts in the tax base. The sample size is 1114.

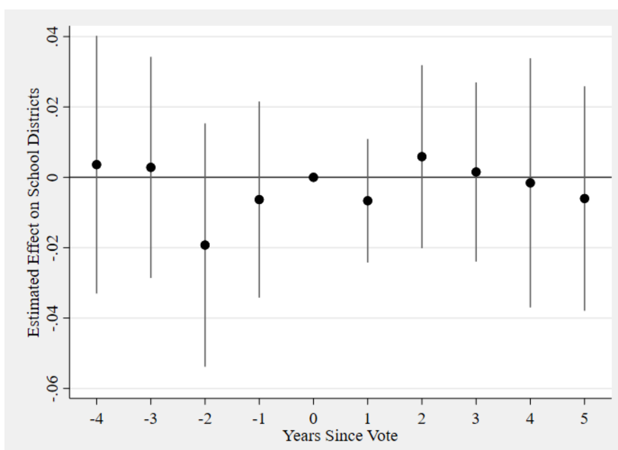


Figure A.2: The Effect of Annexation on K-12 Education

Note: The figure plots the event-study coefficients and 95 percent confidence intervals for the effect of joining the tax base on K-12 expenditure per student using the sample of school districts that are annexed between 1995 and 2011. Standard errors are cluster bootstrapped with 999 replications, clustered at the school district level. The sample size is 985.

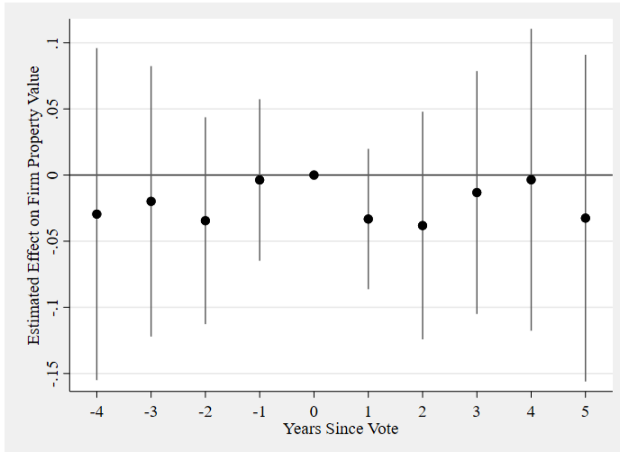
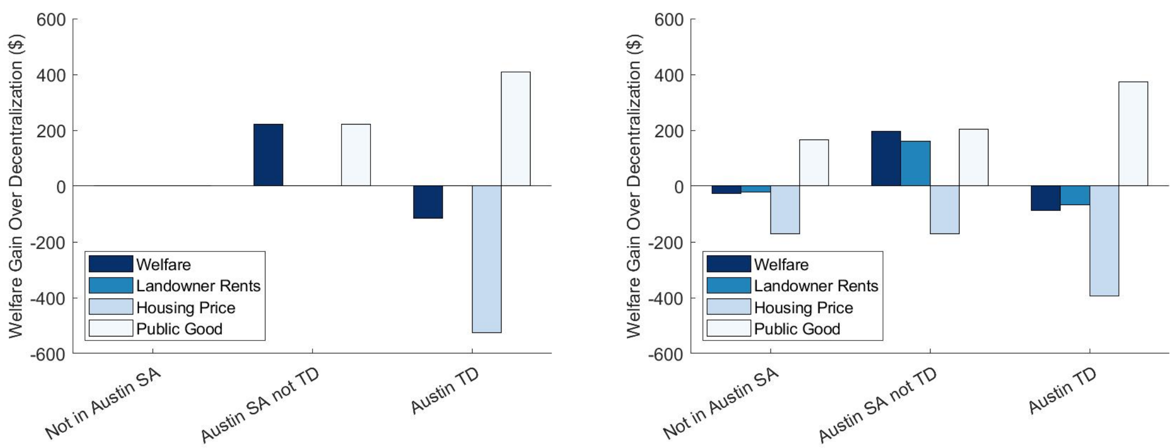


Figure A.3: Effect of Joining the Taxing District on Firms

Note: The solid black circle estimates plot the coefficients of an event-study specification as in Equation 1.1, with the log of firm value as the dependent variable. Year and school district fixed effects are included. The sample size is 1114. Standard errors are cluster bootstrapped with 999 replications, clustered by school district.



(a) Partial Equilibrium

(b) General Equilibrium

Figure A.4: Welfare Decomposition

Note: Each figure is divided into the three community types: those in the current Austin Community College taxing district, those annexed through centralization, and those not in the Austin Community College service area but in the Austin MSA. Panel (a) decomposes the partial equilibrium welfare change when households are immobile and Panel (b) shows the general equilibrium change when they are perfectly immobile for a 10 percent increase in instructional expenditure per student. The left most bar plots the average welfare per capita of baseline residents. The next three bars break down welfare into its components, housing rents to landowners, compensating variation for the housing price change only holding public goods fixed, and compensating variation for the public good change holding the price of housing fixed.

Appendix B

Appendix to Chapter 2

B.1 Optimal Linear Taxes and Minimum Wage

When the federal government has access to linear income taxes, it will always prefer linear taxes to a combination of a minimum wage and linear taxes under the assumptions of our calibrated model.

Consider a binding minimum wage and tax policy $\bar{w}, t \geq 0$. Under the assumptions of our model, mainly a perfectly competitive firm, this set of policy will induce additional unemployment over the competitive equilibrium so that $l(\bar{w}, t)$ low-skilled workers are employed. The minimum wage reduced labor demand while the tax reduced labor supply. The workers earn $\bar{w}$ and high-skilled workers earn $w_h(l(\bar{w}, t))$. The high-skilled wage depends on our structural parameters and the stock of low-skilled employed workers. Now consider $t^\star$ such that $l(0, t^\star) = l(\bar{w}, t)$. It must be the case that $t^\star \geq t$, as a binding minimum wage always reduces low-skilled employment, as long as $l(\bar{w}, t) > 0$. Since l is the same, w_l, w_h, and total output are equal across the policies. However, $t^\star \geq t$ gives that the tax alone raises more revenue. The additional tax revenue is, on net, redistributed to lower consumption agents through the demogrant, increasing total welfare since the social welfare function is concave. As the tax increases, low-skilled labor supply drops, increasing the low-skilled wage and decreasing the high-skilled wage; when it reaches 70%, only 34% of the low-skilled workers are employed, and the low-skilled wage is higher. We may be worried that since the low-skilled face a working cost, they can be high-income but low-consumption and the additional redistribution is harmful for social welfare. However, at that point, income tax collected is redistributed on net to the unemployed from the employed. Since the consumption of the unemployed is always bounded above by the consumption of the employed, even with this tax and demogrant, the tax policy always redistributes to the lower consumption agents.

We need to show that $t^\star$ exists. As long as the low-skilled labor supply is not perfectly inelastic, it will exist because as $l \to 0$, the marginal product goes to infinity, but increasing the tax rate will weakly lower labor supply. When the labor supply is perfectly inelastic (because the low-skilled workers do not face costs from working), then the government can implement a levelling tax, since high-skilled labor supply is also perfectly inelastic to the economy. The value of the levelling tax is maximized when output and labor demand is maximized, i.e. when there is no binding minimum wage.

So the government will always prefer only a linear income tax when both a linear income tax and minimum wage are available in our model.

Appendix C

Appendix to Chapter 3

C.1 Calibration Results

Table C.1 shows that the model matches our target moments well. The college market moments are close to the data, although the private aid elasticities are just outside the boundary of the targeted ranges. The model also approximates the distribution of college attendance by income and ability. Our approach builds on Epple et al. (2017), but can better match these patterns. This is mostly because we allow the no 4-year college option to be a function of both income and ability, rather than fixed across the population. However, our parametrization is highly stylized[1] and the model understates enrollment for the high-income and high-ability. This likely biases our findings against the mechanisms we highlight because it implies that the model understates the value of college quality for the high-ability (weakens the general equilibrium sorting effects), while overstates the value of no college for the high-income (weakens the negative consequences of no longer being admitted to the privates).

C.2 Incidence by Attendance Decision

We present our model implied net benefits and compensating variation for each household of type (b, y, ϵ). Figure C.1 presents the net benefits from the model implied accounting approach conditional on students' enrollment decisions. In panel (a) shows that there are

[1]For tractability, the value of college and utility are continuously differentiable in income and ability, where all households have the same preference parameters (e.g. αs).

three possible values for net benefits based on a household's income. For the very low income who cannot afford the out-of-state public, there are only two possible values. This is because public tuition, net of federal aid, and taxes are determined by income alone. The top line represents those who attend the publics and pay low tuition but receive instructional and capital expenditure, the middle line represents those who attend the privates or do not attend college and therefore only pay taxes, and the bottom line represents those who pay high out-of-state tuition plus taxes in exchange for the public college's services. Although this illustration is helpful for seeing the magnitude of net benefits based on enrollment, it does not show the likelihood of taking up the services by income or ability, which are important for incidence.

Figure C.2 conducts a similar exercise conditional on students' enrollment decisions in both the baseline and counterfactuals using the general equilibrium model and our CV framework. Conditional on college enrollment decisions, CV is mostly linear in income and ability which generates the horizontal lines in the figure. For example, the bottommost line in all panels corresponds to students who attend the out-of-state public in the baseline, but then enroll in the privates when state spending is lower. They are better off because the returns to college quality increased while taxes decreased, and were nearly indifferent between those two options in the baseline. CV for these students is especially low because they no longer help finance public expenditure with high out-of-state tuition. Students who lose access to the highest quality college have much larger CVs and are located above the horizontal lines.

C.3 Incidence Comparative Statics

To further highlight how changes in the private college market affect the incidence of public investment in higher education, we simulate the effect of a 1 percent increase in γ and ω with endogenous sticker prices. Changes in these parameters affect student sorting through the returns to attending a higher quality institution as well as private college admissions decisions and pricing (Equations 3.8 and 3.9). In aggregate, these decisions affect the labor market returns to college as well but the effect is very small for only a 1 percent increase.

For each change in the structural parameters, we simulate the corresponding changes in compensating variation (CV) for a 10 percent decrease in appropriations per student:

$$\begin{aligned} \Delta CV(\gamma) &= CV(1.01\gamma) - CV(\gamma) \\ \Delta CV(\omega) &= CV(1.01\omega) - CV(\omega) \end{aligned} \tag{C.1}$$

and plot the results in Figure C.3. A positive value implies that a household has larger benefits from public investment in higher education when the structural parameter, γ or ω, is high.

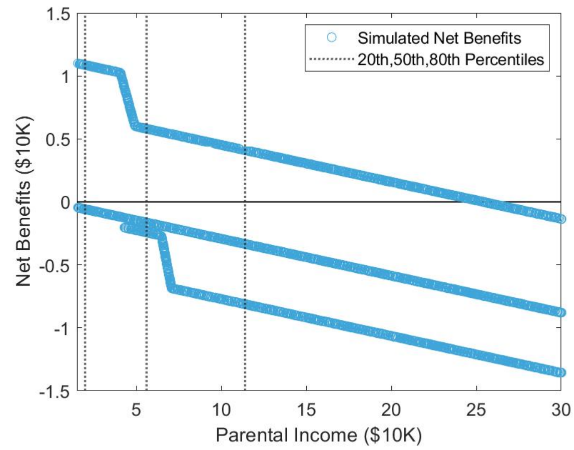
Panels (a) and (b) show how CV changes for a 1 percent increase in γ, or the returns to peer quality, by income and ability respectively. When γ is higher, there is an ambiguous effect on private college access. The implied increase in demand for the highest quality private, because the returns to quality are now higher, lead it to increase prices. Low-income students are less likely to attend and θ_5 falls slightly. At the same time, effective marginal costs are increasing in γ for all students with ability less than the college's average, θ_5. We find that the baseline ($z = \bar{z}$) minimum ability threshold is lower, and the counterfactual threshold rises more after the state spending decrease. Based on these two changes, Panel (b) shows both a large negative and positive spike in $\Delta CV(\gamma)$. When γ is higher, additional students near the 60th percentile of ability are crowded out from decreases in public investment, while fewer are near the 62nd percentile.

Panels (c) and (d) present the effect of a 1 percent increase in ω, or the returns to instructional expenditure, on CV. When the privates value collecting tuition revenue more, additional students with lower ability are admitted to the highest quality private in the baseline with state spending is high. The minimum ability threshold falls, allowing more high-income-modest-ability students to enroll. However, these students are crowded out of the market when state spending falls. More baseline private college students are affected by the decrease in public investment in higher education when the returns to instructional expenditure are high.

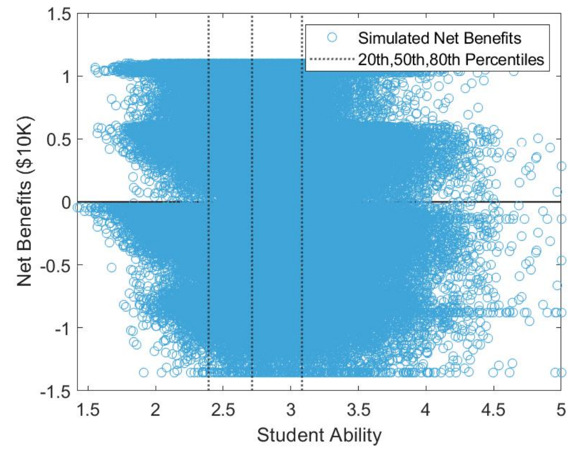
In both comparative statics, others students have a non-zero change in CV largely because the increase in the structural parameters affects the likelihood they enroll in the publics. This creates relatively large changes in public college tuition and relatively small changes in the individual household's value of public spending, ν. Since increases in the structural parameters affect the returns to any college, the returns to college quality, and prices, some students are induced into the publics while others into the private market, generating the horizontal lines in all panels.

Table C.1: Calibration Results

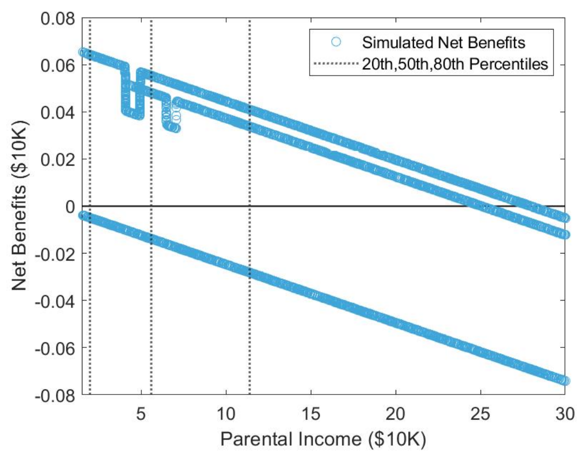
Moment	Data	Model
College Market Moments:		
Average Private Tuition	$27220	$26978
Private Enrollment as Percent of Total	28.5	28.6
Enrollment Elasticity of Aid	0.700	0.700
Additional Private Aid for 1 Std Increase in Ability	(920,1960)	918
Additional Private Aid for $10000 Decrease in Income	(210, 510)	516
Earnings Percent Increase for Marginal-Ability Student	20.0	21.9
Public College Instructional Expenditure	$10463	$10436
Percent Enrolled in College with 2011 Income between:		
$15k and $35k	25.6	21.9
$35k and $55k	32.6	40.2
$55k and $75k	41.2	42.2
$75k and $95k	47.8	49.8
$95k and $135k	54.2	54.9
more than 235k	65.5	62.9
Percent Enrolled in College by Ability quintile:		
First (Lowest)	13.7	12.4
Second	23.4	24.3
Third	32.2	34.2
Fourth	53.9	47.1
Fifth (Highest)	75.9	64.8



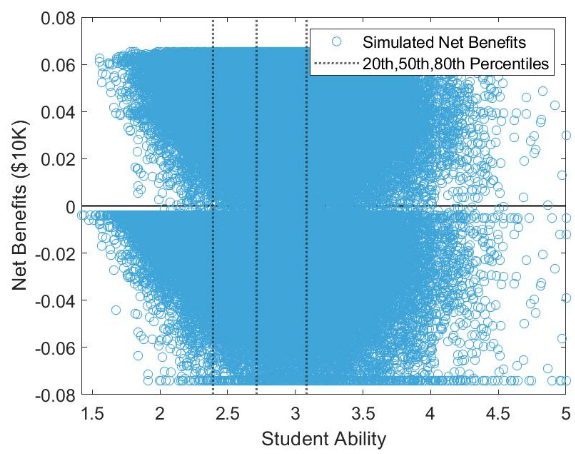
(a) Total net benefits by income



(b) Total net benefits by ability



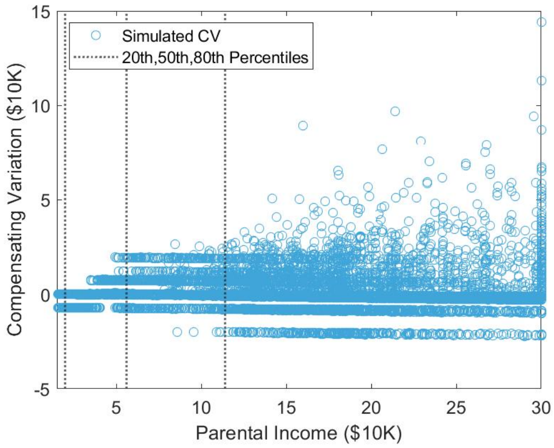
(c) 10% decrease net benefits by income



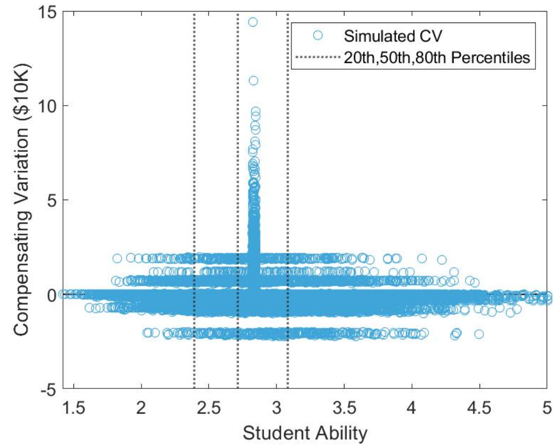
(d) 10% Decrease net benefits by ability

Figure C.1: Net Benefits by Attendance Decisions

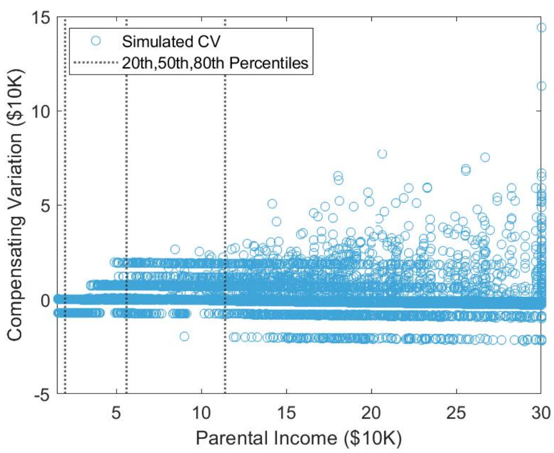
Note: Each open circle represents average net benefits for one draw from the joint distribution of income and ability using the accounting approach based on the household's preference shocks and corresponding enrollment decision. For each of 100,000 households, we take 20 draws from the preference distribution. Panels (a) and (b) present the distribution for public colleges, while (c) and (d) for a 10 percent decrease in appropriations. The dotted vertical lines represent the 20th, 50th and 80th percentiles of the income and ability distribution. A positive value implies that the household of type (b, y, ϵ) receives more in college expenditure than it pays in tuition and taxes.



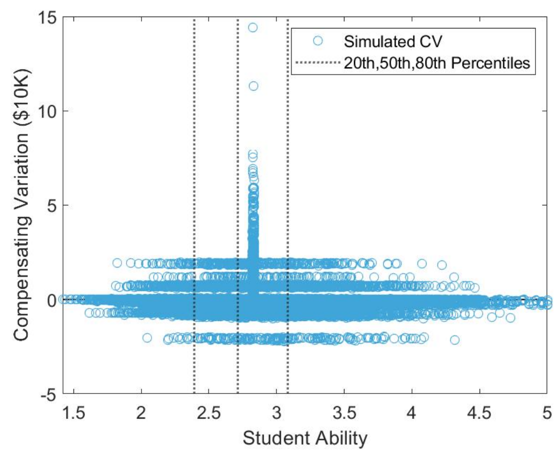
(a) CV by income with fixed sticker prices



(b) CV by ability with fixed sticker prices



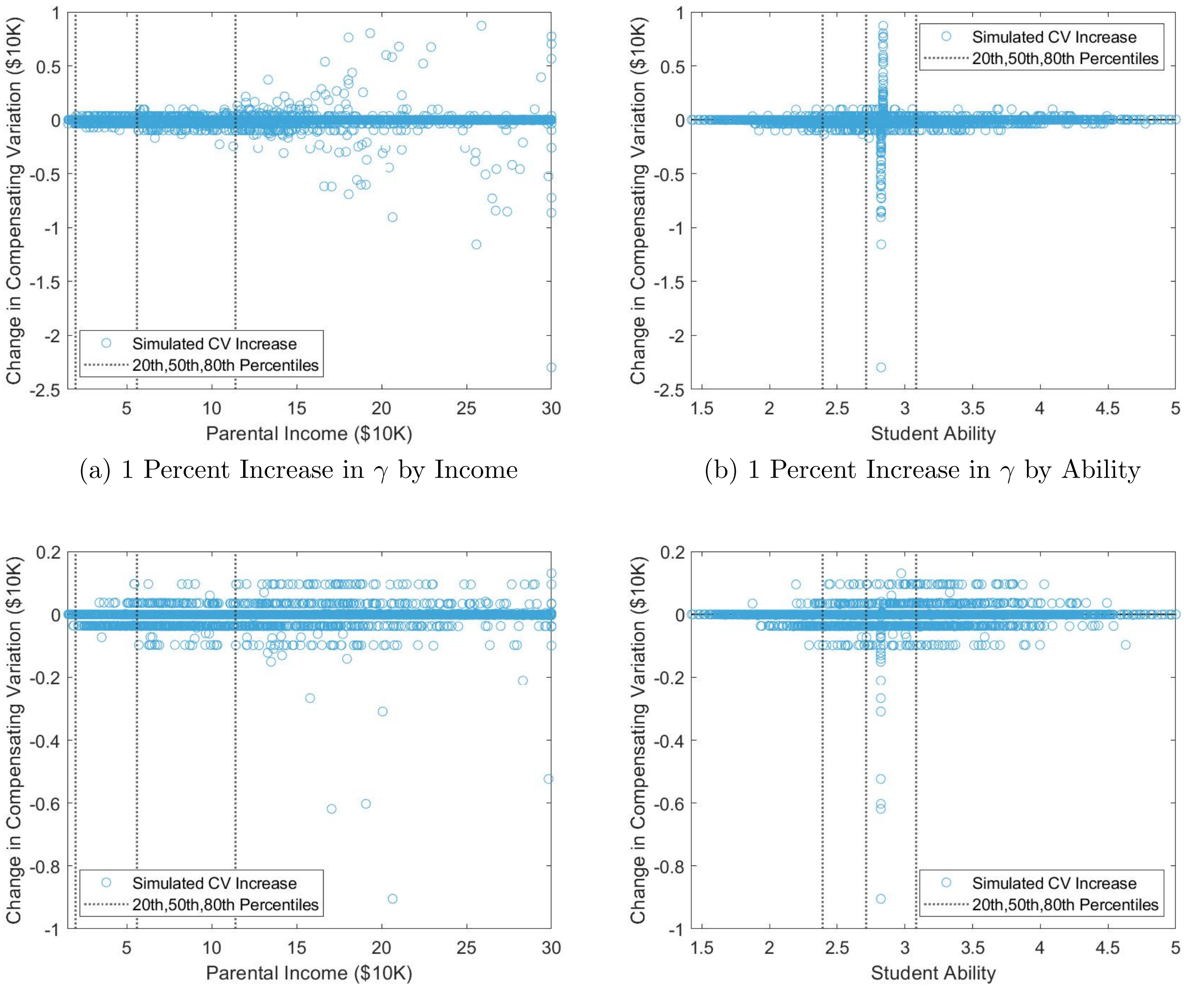
(c) CV by income with endogenous sticker prices



(d) CV by ability with endogenous sticker prices

Figure C.2: Incidence by Attendance Decisions

Note: Each open circle represents average compensating variation (CV) for one draw from the joint distribution of income and ability. For each household, we take 20 sets of draws from the preference distribution, compute the corresponding CV, and then average over the draws. The dotted vertical lines represent the 20th, 50th and 80th percentiles of the income and ability distribution. A positive value in panels (a) and (b) denotes that the household is better off under the baseline than when appropriations decrease by 10 percent, holding private college sticker prices fixed. Panels (c) and (d) compute average CV when private college's sticker prices endogenous adjust.



(a) 1 Percent Increase in γ by Income

(b) 1 Percent Increase in γ by Ability

(c) 1 Percent Increase in ω by Income

(d) 1 Percent Increase in ω by Ability

Figure C.3: Compensating Variation Comparative Statics

Note: Each open circle represents the difference in average compensating variation for one draw from the joint income and ability distribution. For each household, we take 20 sets of draws from the preference distribution, compute the corresponding CV based on the equilibrium changes for the structural parameter and a 1 percent increase in the parameter, and average over the 20 estimates. The dotted vertical lines represent the 20th, 50th, and 80th percentiles of the distributions. A positive value implies that a household has larger benefits from public investments when the parameter (γ or ω) is high.

References

Aaronson, Daniel, Eric French, Isaac Sorkin, and Ted To. "Industry dynamics and the minimum wage: a puttyclay approach." *International Economic Review* 59, no. 1 (2018): 51-84.

Acton, Riley. "Effects of Reduced Community College Tuition on College Choices and Degree Completion." *Education Finance and Policy*, Forthcoming.

Albouy, David, Peter Christensen, and Ignacio Sarmiento-Barbieri. "Unlocking amenities: Estimating public good complementarity." *Journal of Public Economics* 182 (2020): 104110.

Alesina, Alberto, Reza Baqir, and Caroline Hoxby. "Political jurisdictions in heterogeneous communities." *Journal of Political Economy* 112, no. 2 (2004): 348-396.

Allen, Stephen P. "Taxes, redistribution, and the minimum wage: a theoretical analysis." *The Quarterly Journal of Economics* 102, no. 3 (1987): 477-489.

Amior, Michael. "Education and geographical mobility: the role of the job surplus." CEP Discussion Paper, (2019).

Armenter, Roc, and Francesc Ortega. "Credible redistribution policy and skilled migration." *European Economic Review* 55, no. 2 (2011): 228-245.

Anderson, Nick,"Attention, college shoppers. These schools are slashing their prices," The Washington Post, January 21, 2019

Austin Community College District. https://www.austincc.edu/.

Baicker, Katherine, Jeffrey Clemens, and Monica Singhal. "The rise of the states: US fiscal decentralization in the postwar period." *Journal of Public Economics* 96, no. 11-12 (2012): 1079-1091.

Banzhaf, H. Spencer, and B. Andrew Chupp. "Fiscal federalism and interjurisdictional externalities: New results and an application to US Air pollution." *Journal of Public Economics* 96, no. 5-6 (2012): 449-464.

Baum-Snow, Nathaniel, and Fernando Ferreira. "Causal inference in urban and regional economics." In *Handbook of regional and urban economics*, vol. 5, pp. 3-68. Elsevier, 2015.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. "A unified framework for measuring preferences for schools and neighborhoods." *Journal of Political Economy* 115, no. 4 (2007): 588-638.

Bayer, Patrick, Robert McMillan, and Kim Rueben. An equilibrium model of sorting in an urban housing market. No. w10865. *National Bureau of Economic Research*, 2004.

Bednar, Jenna. "The political science of federalism." *Annual Review of Law and Social Science* 7, no. 1 (2011): 269-288

Besley, Timothy, and Stephen Coate. "Centralized versus decentralized provision of local public goods: a political economy approach." *Journal of Public Economics* 87, no. 12 (2003): 2611-2637.

Bettinger, Eric P., Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. "The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment." *The Quarterly Journal of Economics* 127, no. 3 (2012): 1205-1242.

Bloch, Francis, and E. Unal Zenginobuz. "Tiebout equilibria in local public good economies with spillovers." *Journal of Public Economics* 90, no. 8-9 (2006): 1745-1763.

Blom-Hansen, Jens, Kurt Houlberg, Søren Serritzlew, and Daniel Treisman. "Jurisdiction size and local government policy expenditure: Assessing the effect of municipal amalgamation." *American Political Science Review* 110, no. 4 (2016): 812-831.

Blundell, Richard, and Thomas MaCurdy. "Labor supply: A review of alternative approaches." In *Handbook of labor economics*, vol. 3, pp. 1559-1695. Elsevier, 1999.

Boadway, Robin, and Katherine Cuff. "A minimum wage can be welfare-improving and employment-enhancing." *European Economic Review* 45, no. 3 (2001): 553-576.

Bogin, Alexander, William Doerner, and William Larson. "Local house price dynamics: New indices and stylized facts." *Real Estate Economics* 47, no. 2 (2019): 365-398.

Bound, John, Breno Braga, Gaurav Khanna, and Sarah Turner. A passage to America: University funding and international students. No. w22981. *National Bureau of Economic Research*, 2019.

Bound, John, Brad Hershbein, and Bridget Terry Long. "Playing the admissions game: Student reactions to increasing college competition." *Journal of Economic Perspectives* 23, no. 4 (2009): 119-46.

Bricker, Laurie. "Closing the gaps in Texas: The critical role of community colleges." *New Directions for Community Colleges* 2008, no. 141 (2008): 57-65.

Brown, Charles C., and Oates, Wallace E. "Assistance to the Poor in a Federal System." *Journal of Public Economics*, 32 (April 1987): 307-30.

Brueckner, Jan K. "A test for allocative efficiency in the local public sector." *Journal of public Economics* 19, no. 3 (1982): 311-331.

Cadena, Brian C. "Recent immigrants as labor market arbitrageurs: Evidence from the minimum wage." *Journal of Urban Economics* 80 (2014): 1-12.

Calabrese, Stephen, Dennis Epple, Thomas Romer, and Holger Sieg. "Local public good provision: Voting, peer effects, and mobility." *Journal of Public Economics* 90, no. 6-7 (2006): 959-981.

Calabrese, Stephen M., Dennis N. Epple, and Richard E. Romano. "Inefficiencies from metropolitan political and fiscal decentralization: Failures of Tiebout competition." *The Review of Economic Studies* 79, no. 3 (2012): 1081-1111.

Cammenga, Janelle. "Which Places Pay the Most in Property Taxes? *Tax Foundation*, August 13, 2019.

Caplan, Arthur J., Richard C. Cornes, and Emilson CD Silva. "Pure public goods and income redistribution in a federation with decentralized leadership and imperfect labor mobility." *Journal of Public Economics 77*, no. 2 (2000): 265-284.

Chade, Hector, Gregory Lewis, and Lones Smith. "Student portfolios and the college admissions problem." *Review of Economic Studies* 81, no. 3 (2014): 971-1002.

Cheikbossian, Guillaume. "Rent-seeking, spillovers and the benefits of decentralization." *Journal of Urban Economics* 63, no. 1 (2008): 217-228.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. "The effect of minimum wages on low-wage jobs." *The Quarterly Journal of Economics* 134, no. 3 (2019): 1405-1454.

Clemens, Jeffrey, and Michael Wither. "The minimum wage and the Great Recession: Evidence of effects on the employment and income trajectories of low-skilled workers." *Journal of Public Economics* 170 (2019): 53-67.

Colas, Mark and Kevin Hutchinson. "Heterogenous Workers and Federal Income Taxes in a Spatial Equilibrium." *American Economic Journal: Economic Policy*, Forthcoming.

"Consumer Expenditures–2016. U.S. Bureau of Labor Statistics, U.S. Bureau of Labor Statistics, 29 Aug. 2017, www.bls.gov/news.release/cesan.nr0.htm.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst. "Income growth and the distributional effects of urban spatial sorting." No. w26142. *National Bureau of Economic Research*, 2019.

Cremer, Helmuth, and Pierre Pestieau. "Factor mobility and redistribution." In *Handbook of Regional and Urban Economics*, vol. 4, pp. 2529-2560. Elsevier, 2004.

Danziger, Eliav, and Leif Danziger. "The Optimal Graduated Minimum Wage and Social Welfare." In Transitions through the Labor Market: Work, Occupation, Earnings and Retirement, pp. 55-72. Emerald Publishing Limited, 2018.

Denning, Jeffrey T. "College on the cheap: Consequences of community college tuition reductions." *American Economic Journal: Economic Policy* 9, no. 2 (2017): 155-88.

Di Porto, Edoardo, and Sonia Paty. “Cooperation among local governments to deliver public services.” *Politics & Policy* 46, no. 5 (2018): 790-820.

Dixit, Avinash, and John Londregan. ”Fiscal federalism and redistributive politics.” *Journal of Public Economics* 68, no. 2 (1998): 153-180.

Dvorkin, Maximiliano A. “Fiscal Federalism and Optimal Income Taxes.” (2017): 319-336.

Dube, Arindrajit, T. William Lester, and Michael Reich. “Minimum wage shocks, employment flows, and labor market frictions.” *Journal of Labor Economics* 34, no. 3 (2016): 663-704.

Dynarski, Susan, C. J. Libassi, Katherine Michelmore, and Stephanie Owen. Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students. No. w25349. *National Bureau of Economic Research*, 2018.

Economides, George, and Thomas Moutos. “Minimum wages in the Presence of in-kind Redistribution.” (2017).

Epple, Dennis, and Thomas Nechyba. “Fiscal decentralization.” In *Handbook of Regional and Urban Economics*, vol. 4, pp. 2423-2480. Elsevier, 2004.

Epple, Dennis, and Glenn J. Platt. “Equilibrium and local redistribution in an urban economy when households differ in both preferences and incomes.” *Journal of Urban Economics* 43, no. 1 (1998): 23-51.

Epple, Dennis, and Thomas Romer. “Mobility and redistribution.” *Journal of Political Economy* 99, no. 4 (1991): 828-858.

Epple, Dennis, and Richard E. Romano and Holger Sieg. “Admission, tuition, and financial aid policies in the market for higher education.” *Econometrica* 74, no. 4 (2006): 885-928.

Epple, Dennis, Richard Romano, Sinan Sarpça, and Holger Sieg. “A general equilibrium analysis of state and private colleges and access to higher education in the US.” *Journal of Public Economics* 155 (2017): 164-178.

Epple, Dennis, Richard Romano, Sinan Sarpça, and Holger Sieg, and Melanie Zaber. “Market Power and Price Discrimination in the US Market for Higher Education.” *Rand Journal of Economics* (2019)

Epple, Dennis, Thomas Romer, and Holger Sieg. “Interjurisdictional sorting and majority rule: an empirical analysis.” *Econometrica* 69, no. 6 (2001): 1437-1465.

Epple, Dennis, and Holger Sieg. “Estimating equilibrium models of local jurisdictions.” *Journal of Political Economy* 107, no. 4 (1999): 645-681.

Flinn, Christopher J. “Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates.” *Econometrica* 74, no. 4 (2006): 1013-1062.

Ferreyra, Maria Marta. “Estimating the effects of private school vouchers in multidistrict economies.” *American Economic Review* 97, no. 3 (2007): 789-817.

Fillmore, Ian. ”Price Discrimination and Public Policy in the US College Market. Employment Research Newsletter. Volume 23, Number 2.” *WE Upjohn Institute for Employment Research* (2016).

Fu, Chao. “Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market.” *Journal of Political Economy* 122, no. 2 (2014): 225-81.

Fu, Chao, Shoya Ishimaru, and John Kennan. Government Expenditure on the Public Education System. No. w26425. *National Bureau of Economic Research*, 2019.

Fukumura, Koichi, and Atsushi Yamagishi. ”Minimum wage competition.” *International Tax and Public Finance* 27, no. 6 (2020): 1557-1581.

Gerritsen, Aart, and Bas Jacobs. “Is a minimum wage an appropriate instrument for redistribution?.” (2016).

Goodman, Sarena, and Alice Henriques Volz. ”Attendance Spillovers between Public and For-Profit Colleges: Evidence from Statewide Variation in Appropriations for Higher Education.” *Education Finance and Policy* (2016): 1-67.

Goodman-Bacon, Andrew. Difference-in-differences with variation in treatment timing. No. w25018. *National Bureau of Economic Research*, 2018.

Gordon, Grey, and Aaron Hedlund. ”Accounting for the Rise in College Tuition.” Education, Skills, and Technical Change: Implications for Future US GDP Growth (2018): 357-394.

Gould, Elise. “Wages for workers with a high school degree or less rose the fastest over the last year.” Economic Policy Institute. https://www.epi.org/blog/wages-for-workers-with-a-high-school-degree-or-less-rose-the-fastest-over-the-last-year/ (accessed March 1, 2018)

Gordon, Roger H. “An optimal taxation approach to fiscal federalism.” *The Quarterly Journal of Economics* 98, no. 4 (1983): 567-586.

Gordon, Roger H, and Julie Berry Cullen. “Income redistribution in a federal system of governments.” *Journal of Public Economics* 96, no. 11-12 (2012): 1100-1109.

Guesnerie, Roger, and Kevin Roberts. “Minimum wage legislation as a second best policy.” *European Economic Review* 31, no. 1-2 (1987): 490-498.

Hamilton, Jonathan, and Pierre Pestieau. “Optimal income taxation and the ability distribution: Implications for migration equilibria.” *International Tax and public finance* 12, no. 1 (2005): 29-45.

Hansen, W. Lee, and Burton A. Weisbrod. “The distribution of costs and direct benefits of public higher education: the case of California.” *Journal of Human Resources* (1969): 176-191.

Hansen, W. Lee. “Income Distribution Effects of Higher Education.” The American Economic Review 60, no. 2 (1970): 335-40.

Hoekstra, Mark. ”The effect of attending the flagship state university on earnings: A discontinuity-based approach." *The Review of Economics and Statistics* 91, no. 4 (2009): 717-724.

Holmes, Thomas J., and Holger Sieg. "Structural estimation in urban economics." In *Handbook of regional and urban economics*, vol. 5, pp. 69-114. Elsevier, 2015.

Hoxby, Caroline, and Christopher Avery. "The missing" one-offs": The hidden supply of high-achieving, low-income students." *Brookings Papers on Economic Activity* 2013.1 (2013): 1-65.

Hoxby, Caroline, and Sarah Turner. "What high-achieving low-income students know about college." *American Economic Review* 105.5 (2015): 514-17.

Hoynes, Hilary W., and Erzo FP Luttmer. "The insurance value of state tax-and-transfer programs." *Journal of Public Economics* 96, no. 11-12 (2012): 1110-1128.

Hungerbühler, Mathias, and Etienne Lehmann. "On the optimality of a minimum wage: New insights from optimal tax theory." *Journal of Public Economics* 93, no. 3-4 (2009): 464-481.

Janeba, Eckhard, and John Douglas Wilson. "Optimal fiscal federalism in the presence of tax competition." *Journal of Public Economics* 95, no. 11-12 (2011): 1302-1311.

Johnson, George E. "Subsidies for Higher Education." *Journal of Labor Economics* 2, no. 3 (1984): 303-18.

Johnson, William R. "Are public subsidies to higher education regressive?" *Education Finance and Policy* 1.3 (2006): 288-315.

Kasman, Matt, and Katherine Guyot. "Predicting the Impact of College Subsidy Programs on College Enrollment." Brookings. October 21, 2019.

Katz, Lawrence F., and Kevin M. Murphy. "Changes in relative wages, 1963 – 1987: supply and demand factors." *The Quarterly Journal of Economics* 107, no. 1 (1992): 35-78.

Keen, Michael J., and Christos Kotsogiannis. "Does federalism lead to excessively high taxes?." *American Economic Review* 92, no. 1 (2002): 363-370.

Keen, Michael J., and Christos Kotsogiannis. "Tax competition in federations and the welfare consequences of decentralization." *Journal of Urban Economics* 56, no. 3 (2004): 397-407.

Kennan, John, and James R. Walker. "The effect of expected income on individual migration decisions." *Econometrica* 79, no. 1 (2011): 211-251.

Kleven, Henrik, Camille Landais, Mathilde Muoz, and Stefanie Stantcheva. Taxation and Migration: Evidence and Policy Implications. No. w25740. National Bureau of Economic Research, 2019.

Kline, Patrick, and Enrico Moretti. "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs." *Annual Review of Economics* 6, no. 1 (2014): 629-662.

Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins. “The new economics of equilibrium sorting and policy evaluation using housing markets.” *Journal of economic literature* 51, no. 4 (2013): 1007-62.

“Labor Force Statistics from the Current Population Survey.” Bureau of Labor Statistics. https://www.bls.gov/cps/ (accessed March 1, 2018)

Ladd, Helen F., and Doolittle, Fred C. “Which Level of Government Should Assist the Poor?” *National Tax Journal*, 35 (September 1982): 323-36.

Lavecchia, Adam. “Minimum Wage Policy with Optimal Taxes and Unemployment.” McMaster University Department of Economics Working Paper 2019-03 (2019).

Lee, David, and Emmanuel Saez. “Optimal minimum wage policy in competitive labor markets.” *Journal of Public Economics* 96, no. 9-10 (2012): 739-749.

Lochner, Lance, and Alexander Monge-Naranjo. ”Credit constraints in education.” *Annual Review of Economics* 4, no. 1 (2012): 225-256.

Lockwood, Ben. ”Distributive politics and the costs of centralization.” *The Review of Economic Studies* 69, no. 2 (2002): 313-337.

Luttmer, Erzo FP. ”Does the minimum wage cause inefficient rationing?.” *The BE Journal of Economic Analysis & Policy* 7, no. 1 (2007).

Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles. IPUMS: Version 14.0. Minneapolis, MN: IPUMS. 2019. http://doi.org/10.18128/D050.V14.0

Mellow, Gail O., and Cynthia M. Heelan. *Minding the dream: The process and practice of the American community college*. Rowman & Littlefield, 2014.

McDuff, DeForest. ”Quality, tuition, and applications to in-state public colleges.” *Economics of Education Review* 26, no. 4 (2007): 433-449.

McFadden, Daniel. ”The measurement of urban travel demand.” *Journal of Public Economics* 3, no. 4 (1974): 303-328.

McFarlin Jr, Isaac, Paco Martorell, and Brian McCall. “Do public subsidies improve college attainment and labor market outcomes? Evidence from community college taxing district expansions.” (2017).

McFarlin Jr, Isaac, Brian McCall, and Paco Martorell. “How much do tuition subsidies promote college access? Evidence from community college taxing districts.” (2018).

Moretti, Enrico. ”Human capital externalities in cities.” In *Handbook of regional and urban economics*, vol. 4, pp. 2243-2291. Elsevier, 2004.

Moretti, Enrico. ”Real wage inequality.” *American Economic Journal: Applied Economics* 5, no. 1 (2013): 65-103.

Mountjoy, Jack. ”Community Colleges and Upward Mobility.” (2019).

Neumark, David. “The Econometrics and Economics of the Employment Effects of Minimum Wages: Getting from Known Unknowns to Known Knowns.” NBER Working Paper, (2018).

Neumark, David, and William L. Wascher. ”Minimum wages and employment.” Foundations and Trends in Microeconomics 3, no. 12 (2007): 1-182.

Oates, Wallace E. “The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the Tiebout hypothesis.” *Journal of Political Economy* 77, no. 6 (1969): 957-971.

Oates, Wallace E. “Fiscal Federalism.” (1972).

Oates, Wallace E. “An economist perspective on Fiscal federalism-in: The Political Economy of Fiscal Federalism (ed. by. WE Oates), 1977.” (1977): 3-19.

Pechman, Joseph A. “The Distributional Effects of Public Higher Education in California”. *Journal of Human Resources* 5 (1971): 361-370.

Pechman, Joseph A. “The Distribution of Costs and Benefits of Public Higher Education: Further Comments”. *Journal of Human Resources* 6 (1971): 375-3

Phillips, David C. ”Do low-wage employers discriminate against applicants with long commutes? Evidence from a correspondence experiment.” *Journal of Human Resources* (2018): 1016-8327R.

Poterba, James M. “Taxation and Housing: Old Questions, New Answers.” *American Economic Review* 82, no. 2 (1992): 237-242.

Rizzo, Michael, and Ronald G. Ehrenberg. “Resident and nonresident tuition and enrollment at flagship state universities.” College choices: The economics of where to go, when to go, and how to pay for it. University of Chicago Press, 2004. 303-354.

Rothschild, Michael, and Lawrence J. White. “The analytics of the pricing of higher education and other services in which the customers are inputs.” *Journal of Political Economy* 103, no. 3 (1995): 573-586.

Roback, Jennifer. ”Wages, rents, and the quality of life.” *Journal of political Economy* 90, no. 6 (1982): 1257-1278.

Rohlin, Shawn M. ”State minimum wages and business location: Evidence from a refined border approach.” *Journal of Urban Economics* 69, no. 1 (2011): 103-117.

Rosen, Sherwin. (1979). Wage-based indexes of urban quality of life. Baltimore: Johns Hopkins University.

Ross, Stephen, and John Yinger. “Sorting and voting: A review of the literature on urban public finance.” In *Handbook of regional and urban economics* 3 (1999): 2001-2060.

Saarimaa, Tuukka, and Janne Tukiainen. “I don’t care to belong to any club that will have me as a member: empirical analysis of municipal mergers.” *Political Science Research and*

Methods 2, no. 1 (2014): 97-117.

Saiz, Albert. "The geographic determinants of housing supply." *The Quarterly Journal of Economics* 125, no. 3 (2010): 1253-1296.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh. "Interjurisdictional housing prices in locational equilibrium." *Journal of Urban Economics* 52, no. 1 (2002): 131-153.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh. "Estimating the general equilibrium benefits of large changes in spatially delineated public goods." *International Economic Review* 45, no. 4 (2004): 1047-1077.

Simon, Andrew and Matthew Wilson. "Optimal Minimum Wage Setting in a Federal System." *Journal of Urban Economics*, Forthcoming.

Simon, Andrew. "Public Good Spillovers and Fiscal Centralization: Evidence from Community College Expansions." (2020).

Smith, V. Kerry, Holger Sieg, H. Spencer Banzhaf, and Randall P. Walsh. "General equilibrium benefits for environmental improvements: projected ozone reductions under EPA's Prospective Analysis for the Los Angeles air basin." *Journal of Environmental Economics and Management* 47, no. 3 (2004): 559-584.

Stigler, George J. "The Tenable Range of Functions of Local Government." In Federal Expenditure Policy for Economic Growth and Stability, by the Joiny Economic Committee. Washington: Government Printing Office, 1957.

Stiglitz, Joseph E. "Self-selection and Pareto efficient taxation." *Journal of Public Economics* 17, no. 2 (1982): 213-240.

Surez Serrato, Juan Carlos, and Owen Zidar. "Who benefits from state corporate tax cuts? A local labor markets approach with heterogeneous firms." *American Economic Review* 106, no. 9 (2016): 2582-2624.

Texas Association of Community Colleges. https://tacc.org/tacc/texas-community-colleges.

Texas Higher Education Coordinating Board. http://www.thecb.state.tx.us.

Tiebout, Charles M. "A pure theory of local expenditures." *Journal of Political Economy* 64, no. 5 (1956): 416-424.

Tricaud, Clemence. Better alone? evidence on the costs of intermunicipal cooperation. CREST Working paper, 2019.

United States. Bureau of Justice Statistics. Law Enforcement Agency Identifiers Crosswalk, United States, 2012. Inter-university Consortium for Political and Social Research, 2018-09-18. https://doi.org/10.3886/ICPSR35158.v2

Urban Institute. "Property Taxes. July 15, 2019.

Van Der Klaauw, Wilbert. "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach." *International Economic Review* 43, no. 4 (2002): 1249-287.

Walsh, Randy. "Endogenous open space amenities in a locational equilibrium." *Journal of Urban Economics* 61, no. 2 (2007): 319-344.

Webber, Douglas A. "State divestment and tuition at public institutions." *Economics of Education Review* 60 (2017): 1-4.

Weese, Eric. "Political mergers as coalition formation: An analysis of the Heisei municipal amalgamations." '*Quantitative Economics* 6, no. 2 (2015): 257-307.

Wildasin, David E. "Labor-market integration, investment in risky human capital, and fiscal competition." *American Economic Review* 90.1 (2000): 73-95.

Wildasin, David E. "Income distribution and redistribution within federations." *Annales d'Economie et de Statistique* (1997): 291-313.

Wilson, John D. "A theory of interregional tax competition." *Journal of Urban Economics* 19, no. 3 (1986): 296-315.

Wildasin, David E. "Trade, Capital Mobility, and Tax Competition." *Journal of Political Economy.* 95 (August 1987): 835-56.

Wildasin, David E. "Tax competition in a federal setting." *Handbook of Multilevel Finance* (2015): 264.

Wildasin, David E. "Optimal linear income taxation in the presence of emigration." *Journal of Public Economics* 18, no. 3 (1982): 363-379.

Zodrow, George R., and Peter Mieszkowski. "Pigou, Tiebout, property taxation, and the underprovision of local public goods." *Journal of Urban Economics* 19, no. 3 (1986): 356-370.

Windham, Douglas M. Education, Equality and Income Redistribution. (1970) Lexington: D.C. Heath.

Zimmerman, Seth D. "The returns to college admission for academically marginal students." *Journal of Labor Economics* 32, no. 4 (2014): 711-754.